

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 18:07:04 ; Search time 67 Seconds
(without alignments)
421.629 Million cell updates/sec

Title: US-09-817-198A-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDGVG.....LEEEGKPEGPANSSKTWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_l01002.*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1105	100.0	401	AAU17136	Novel signal trans
2	1092	98.8	218	AAAB41604	Human OREF1368
3	832	75.3	188	AAU17555	Novel signal trans
4	546.5	49.5	204	ABB70670	Drosophila melanog
5	545	49.3	213	ABB11916	Human rab8 homolog
6	540	48.9	221	ABP41333	Human ovarian anti
7	532	48.1	246	AAAB58196	Lung cancer associ
8	530.5	48.0	200	AAAB19165	Amino acid sequenc
9	530.5	48.0	200	AAAB09979	Human Rab10 protei
10	530.5	48.0	200	AAAB95340	Human protein sequ

11	530.5	48.0	218	21	AAAB56993	Human prostate can
12	529.5	47.9	218	22	ABG07266	Novel human diagno
13	528.5	47.8	207	22	ABB71647	Drosophila melanog
14	528.5	47.8	207	22	ABG67154	Amino acid sequenc
15	528.5	47.8	207	22	AAAB94628	Human protein sequ
16	526	47.6	201	21	AAAB09982	Canine Rab10 protei
17	523	47.3	199	21	AAAB09980	Human Rab10 protei
18	523	47.3	199	21	AAAB09981	Human Rab10 protei
19	519	47.0	216	21	AAAG08688	Arabidopsis thalia
20	519	47.0	216	21	AAAG53945	Arabidopsis thalia
21	519	47.0	253	21	AAAG53944	Arabidopsis thalia
22	519	47.0	254	21	AAAG08687	Arabidopsis thalia
23	513.5	46.5	215	21	AAAG35215	Zea mays protein f
24	510	46.2	218	21	AAAG19220	Arabidopsis thalia
25	507	45.9	216	21	AAAG08006	Arabidopsis thalia
26	506.5	45.8	209	22	ABG23365	Novel human diagno
27	501	45.3	224	21	ABG47826	Arabidopsis thalia
28	501	45.3	234	21	AAAG47825	Arabidopsis thalia
29	501	45.3	335	22	ABG23366	Novel human diagno
30	478.5	43.3	201	20	AAAY00919	Human Rab protein,
31	478.5	43.3	201	22	AAU28024	Novel human secret
32	478.5	43.3	221	22	AAO13525	Human polypeptide
33	478.5	43.3	224	21	AAAB58758	Breast and ovarian
34	476	43.1	203	21	AAAG30499	Arabidopsis thalia
35	476	43.1	258	21	AAAG30498	Arabidopsis thalia
36	468	42.4	202	21	AAAG07763	Arabidopsis thalia
37	467	42.3	205	21	AAAB34843	Gene 44 human secr
38	467	42.3	205	21	AAAB34844	Human secreted pro
39	466	42.2	202	21	AAAG10858	Arabidopsis thalia
40	462	41.8	205	22	ABBS9808	Drosophila melanog
41	457	41.4	218	21	AAAG30710	Arabidopsis thalia
42	457	41.4	221	21	AAAG30709	Arabidopsis thalia
43	457	41.4	254	21	AAAG45323	Arabidopsis thalia
44	456.5	41.3	227	21	AAAB19163	Amino acid sequenc
45	456.5	41.3	227	22	AAAB39600	Human polypeptide

ALIGNMENTS

RESULT 1

AAU17136
ID AAU17136 standard; Protein; 401 AA.

XX
AC AAU17136;

XX
DT 07-NOV-2001 (first entry)

DE
XX Novel signal transduction pathway protein, Seq ID 701.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antinflammatory; anti-HIV; antibacterial; antinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disease; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; AIDS;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.

XX
OS Homo sapiens.

XX
PN WO200154733-A1.

XX
PD 02-AUG-2001.

XX
PF 17-JAN-2001; 2001WO-US01312.

XX
PR 31-JAN-2000; 2000US-0179065.

PR
04-FEB-2000; 2000US-0180628.

PR
24-FEB-2000; 2000US-0184664.

PR
02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231415.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.
XX
XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465460/50.

N-PSDB; AAS27053.

Novel polypeptides useful for diagnosing, treating, preventing and/or
prognosing disorders related to the proteins, including cancers, immune
disorders and neuronal disorders -

Claim 1; SEQ ID No 701; 880pp; English.

XX

CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.

Query Match 100.0%; Score 1105; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.1e-108;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLLLIGDSGVGKTCCLLCRFDTDFNEFSSHSHTIGVDFKMKTIYVDGKIVRTQ 60
Db 27 MAKQYDVLFRLLLLIGDSGVGKTCCLLCRFDTDFNEFSSHSHTIGVDFKMKTIYVDGKIVRTQ 86
QY 61 IWDTAGOERYQITTKQYRRAGQIFLVYDISSERSYOHIMKMWSDVDYAPGVQKILIG 120
Db 87 IWDTAGOERYQITTKQYRRAGQIFLVYDISSERSYOHIMKMWSDVDYAPGVQKILIG 146
QY 121 NKADEEQKRVQREGQOQQLAKYGMDFYETSACTNLNLIKESFTRLTVELVLAHRKEGL 180
Db 147 NKADEEQKRVQREGQOQQLAKYGMDFYETSACTNLNLIKESFTRLTVELVLAHRKEGL 206
QY 181 RMRASNELALAELEEEGKPEGPNSSKTCWC 212
Db 207 RMRASNELALAELEEEGKPEGPNSSKTCWC 238

RESULT 2
AAB41604

ID AAB41604 standard; Protein; 218 AA.

XX AAB41604;

XX 08-FEB-2001 (first entry)

DE Human ORF1368 polypeptide sequence SEQ ID NO:2736.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

OS Homo sapiens.

XX ID AAU17555 standard; Protein; 188 AA.

XX AAU17555;

AC AAU17555;

XX 05-OCT-2000.
PD 31-MAR-2000; 2000WO-US08621.
PF 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
PA Shinkets RA, Leach M;
PI WPI; 2000-602362/57.
XX N-PSDB; AAC75813.

XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -

PS Claim 11; Page 1979-1980; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX SQ Sequence 218 AA;

Query Match 98.8%; Score 1092; DB 21; Length 218;

Best Local Similarity 97.2%; Pred. No. 2e-107;

Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 MAKQYDVLFRLLLLIGDSGVGKTCCLLCRFDTDFNEFSSHSHTIGVDFKMKTIYVDGKIVRTQ 60
Db 1 MAKQYDVLFRLLLLIGDSGVGKTCCLLCRFDTDFNEFSSHSHTIGVDFKMKTIYVDGKIVRTQ 60
QY 61 IWDTAGOERYQITTKQYRRAGQIFLVYDISSERSYOHIMKMWSDVDYAPGVQKILIG 120
Db 61 IWDTAGOERYQITTKQYRRAGQIFLVYDISSERSYOHIMKMWSDVDYAPGVQKILIG 120
QY 121 NKADEEQKRVQREGQOQ-----QLAKYGMDFYETSACTNLNLIKESFTRLTVELVLAHR 174
Db 121 NKADEEQKRVQREGQOQKPSQLAKYGMDFYETSACTNLNLIKESFTRLTVELVLAHR 180
QY 175 KELEGLMRASNELALAELEEEGKPEGPNSSKTCWC 212
Db 181 KELEGLMRASNELALAELEEEGKPEGPNSSKTCWC 218

RESULT 3

AAU17555

ID AAU17555 standard; Protein; 188 AA.

XX AAU17555;

AC AAU17555;

XX 07-NOV-2001 (first entry)
DT Novel signal transduction pathway protein, Seq ID 1120.
DE
XX
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
XX acquired immune deficiency syndrome.
OS Homo sapiens.
XX
XX WO200154733-A1.
PN
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01312.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231415.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.


```
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-463460/50.
DR N-PSDB; AAS27472.
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
XX Claim 1; SEQ ID NO 1120; 880pp; English.
XX
XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX

Query Match 75.3%; Score 832; DB 22; Length 188;
Best Local Similarity 97.6%; Pred. No. 6.2e-80;
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDVKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDVKMTIEVDGKVRQ 83
QY 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISSERSYQHIMKWSDVDEYAPGVQKILIG 120
DB 84 IWDTAGQERYQTITKQYRRAGGIFLVYDISSERSYQHIMKWSDVDEYAPGVQXILIG 143
QY 121 NKADEEQKRVQGGQOLAKEYGMDFYETSACTNLNLIKESFTR 164
DB 144 NKADEEQKRVQGGQOLAXEYGMDFYETSACTNLNLIKESFTR 187

RESULT 4
ABB70670
ID ABB70670 standard; Protein: 204 AA.
XX

AC ABB70670;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 38802.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL14773.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 38802; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 204 AA;
XX
XX Query Match 49.5%; Score 546.5; DB 22; Length 204;
Best Local Similarity 49.8%; Pred. No. 1.4e-49;
Matches 105; Conservative 47; Mismatches 50; Indels 9; Gaps 3;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDVKMTIEVDGKVRQ 59
DB 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDVKMTIEVDGKVRQ 60
QY 60 QIWDTAGQERYQTITKQYRRAGGIFLVYDISSERSYQHIMKWSDVDEYAPGVQKILI 119
DB 61 QIWDTAGQERYQTITTSYRGAMGIMLVYDITNEKSFENIVKWLNRNIDEHANEDVEKMIL 120
QY 120 GNKADEEQKRVQGGQOLAKEYGMDFYETSACTNLNLIKESFTRTELTLVLQAHKELEG 179
DB 121 GNKCDMTKRVVKNRGERGAIAHREHIGREHIGREHIGREHIGREHIGREHIGREHIG 176
QY 180 LMRASNLALAELEEECKPEGPANSSKTC 210
DB 177 RESAENQERVIIDRRNQEKAP----GYSKCC 203

RESULT 5
ABB11916
ID ABB11916 standard; peptide: 213 AA.
XX
XX ABB11916;
AC
```


XX PS Claim 11; SEQ ID No 2465; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP4328) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 221 AA;

Query Match 48.9%; Score 540; DB 23; Length 221;
Best Local Similarity 52.7%; Pred. No. 7.5e-49;
Matches 98; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Qy 2 AKOYDVLFRLLIGDSGVGKTCCLRCFTDNEFHSHSHISTIGVDFKMTIEVDGKIVRIQI 61
Db 16 AKTYDLFRLLIGDSGVGKTCCLRCFTDNEFHSHSHISTIGVDFKMTIEVDGKIVRIQI 75

Qy 62 WDTAGQERYQITTKOYYRRAQGIPLVYDIDSSERSYQHIMKMWSDVDEYAPGVQKILGN 121
Db 76 WDTAGQERYQITTKOYYRRAQGIPLVYDIDSSERSYQHIMKMWSDVDEYAPGVQKILGN 135

Qy 122 KADEEQKRVGREGQQQLAKEYGMDFYETSACTNLNKESTRLTELVLQAHKRELEGLR 181
Db 136 KCDVNDKRVGREGQQQLAKEYGMDFYETSACTNLNKESTRLTELVLQAHKRELEGLR 195

Qy 182 MRASNE 187
Db 196 PQGSNQ 201

RESULT 7
AAB58196
ID AAB58196 standard; Protein; 246 AA.

XX AC AAB58196;
XX DT 14-MAR-2001 (first entry)
XX DE Lung cancer associated polypeptide sequence SEQ ID 534.
XX KW Human; lung cancer associated protein; neuroprotective; cytostatic;
KW KW cardioactive; immunomodulatory; muscular active; vulnerable;
KW KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW KW proliferative disorder; wound healing; infectious disease.
XX OS Homo sapiens.

XX PN WO200055180-A2.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05918.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.
FA (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI: 2000-587514/55.
N-PSDB; AAF18072.

XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -

XX Claim 11; Page 1027-1028; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnerable; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.

XX SQ Sequence 246 AA;

Query Match 48.1%; Score 532; DB 21; Length 246;
Best Local Similarity 47.6%; Pred. No. 6.2e-48;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLRCFTDNEFHSHSHISTIGVDFKMTIEVDGKIVRIQ 60
Db 44 MAKAYDHLFKLLIGDSGVGKTCCLRIIRFAEDNFNNTYISTIGIDFKIRTVIEGKKIKLQ 103

Qy 61 IWDTAGQERYQITTKOYYRRAQGIPLVYDIDSSERSYQHIMKMWSDVDEYAPGVQKILIG 120
Db 104 VMDTAGQERYQITTKOYYRRAQGIPLVYDIDSSERSYQHIMKMWSDVDEYAPGVQKILIG 163

Qy 121 NKADEEQKRVGREGQQQLAKEYGMDFYETSACTNLNKESTRLTELVLQAHKRELEGLR 180
Db 164 NKCDMEARKVKQEQADKLAREHGIRFETSAKSMNVNDEAFSLARDIL-----LKSG 217

Qy 181 RMASNELALAELEEEGKPGEPANSSKTC 210
Db 218 GRRSGN-----GNKP--PSTD LKTC 235

RESULT 8
AAB19165
ID AAB19165 standard; Protein; 200 AA.

XX AC AAB19165;
XX DT 19-FEB-2001 (first entry)
XX OS Homo sapiens.

DE Amino acid sequence of human RAB10.
KW RAB protein; GTPase; GTP binding; gene therapy; cancer.
XX
OS Homo sapiens.
XX
PN WO200058464-A2.
XX
PD 05-OCT-2000.
XX
PF 13-MAR-2000; 2000WO-US06330.
XX
PR 25-MAR-1999; 99US-0126083.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;
XX
DR WPI; 2000-647233/62.
DR N-PSDB; AAA96887.
XX
XX Novel isolated nucleic acid encoding a mammalian RAB protein useful for
PT identifying homologous or related genes, in producing composition that
PT modulates expression or function of RAB for cancer therapy -
XX
PS Claim 2; Page 39; 58pp; English.
XX
CC The present sequence represents a mammalian RAB protein. RAB proteins
CC constitute the largest family of small GTPases, with over 40 identified
CC isoforms. RAB proteins contain four highly conserved peptide sequences
CC involved in GTP binding and hydrolysis. Compositions comprising RAB
CC nucleic acid are useful for identifying homologous or related genes,
CC in producing compositions that modulate the expression or function of
CC RAB, for gene therapy, mapping functional regions of the protein and
CC in studying associated physiological pathways. In addition, modulation
CC of the gene activity in vivo is used for prophylactic and therapeutic
CC purposes, such as treatment of cancer, and identification of cell type
CC based on expression. The DNA may also be used to identify expression of
CC the gene in a biological specimen.
XX
XS Sequence 200 AA;
XX
Query Match 48.0%; Score 530.5; DB 21; Length 200;
Best Local Similarity 55.8%; Pred. No. 6.6e-48;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKO-YDVLFRLLIGDSGVGKTCVLRFDNEFHSSTIGVDFKMTIEVDGIKVI 59
Db 1 MAKTYDLLFKLLIGDSGVGKTCVLRFDNEFHSSTIGVDFKMTIEVDGIKVI 60
QY 60 QIWDTAGQERFYQITTKOYRRAGIPLVYDISSERSYQHIMKWSVDVYAPGVQKILI 119
Db 61 QIWDTAGQERFHTTTSYRGAMGIMLVYDITNGKSFENISKWLRNIDEHANEDVERMLL 120
QY 120 GNKADEQKRGVREGQQLAKYGMDFYETSACTNLNIKESFTRLTELVLQ 171
Db 121 GNKCDMDKRVKPGKEQIAREHGIRFFETSAKANINIEKAFLLTAEIDL 172
RESULT 9
AAB09979
ID AAB09979 standard; Protein; 200 AA.
XX
AC AAB09979;
XX
DT 19-OCT-2000 (first entry)
XX
DE Human Rab10 protein.
XX
KW Human; Rab10; Yap/Rab family.
XX
OS Homo sapiens.
XX

PN CN1249345-A.
XX
PD 05-APR-2000.
XX
PE 28-SEP-1998; 98CN-0121911.
XX
PR 28-SEP-1998; 98CN-0121911.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Tu Q, Gao J;
XX
DR WPI; 2000-400723/35.
DR N-PSDB; AAA40104.
XX
PT Preparation of human gene coding sequence, its encoded polypeptide -
XX
PS Claim 2; Page 18; 23pp; Chinese.
XX
CC This invention describes a novel human Rab10 cDNA sequence. The protein
CC coded by said sequence belongs to Yap/Rab family and is the homolog of
CC mouse Rab10. The present invention also relates to the polypeptide coded
CC by said nucleotide sequence and the application and preparing process
CC of said polynucleotide and said polypeptide. This sequence represents the
CC human Rab10 protein described in the method of the invention.
XX
XS Sequence 200 AA;
XX
Query Match 48.0%; Score 530.5; DB 21; Length 200;
Best Local Similarity 55.8%; Pred. No. 6.6e-48;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKO-YDVLFRLLIGDSGVGKTCVLRFDNEFHSSTIGVDFKMTIEVDGIKVI 59
Db 1 MAKTYDLLFKLLIGDSGVGKTCVLRFDNEFHSSTIGVDFKMTIEVDGIKVI 60
QY 60 QIWDTAGQERFYQITTKOYRRAGIPLVYDISSERSYQHIMKWSVDVYAPGVQKILI 119
Db 61 QIWDTAGQERFHTTTSYRGAMGIMLVYDITNGKSFENISKWLRNIDEHANEDVERMLL 120
QY 120 GNKADEQKRGVREGQQLAKYGMDFYETSACTNLNIKESFTRLTELVLQ 171
Db 121 GNKCDMDKRVKPGKEQIAREHGIRFFETSAKANINIEKAFLLTAEIDL 172
RESULT 10
AAB95340
ID AAB95340 standard; Protein; 200 AA.
XX
AC AAB95340;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:17619.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX

PD 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
PF 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
PR (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS71453.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 37625; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABB00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 218 AA;
SQ Query Match 47.9%; Score 529.5; DB 22; Length 218;
Best Local Similarity 55.8%; Pred. No. 9.6e-48;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
Qy 1 MAKO-YDVLFRLLIGSGVGKTCCLCRFTDNEPHSSHISTIGYDFKMKTIKIVGKIRI 59
Db 19 MAKRTYDLLEKLLIGSGVGKTCVLFPSDDAFNTFTSTIGIDFKIRIETLKNKIKLQ 78
Qy 60 QIWTAGQERYOTTTKYIRRAQGIIFLVYDITSSERSYQHIMKWYSDVDEVAPEGVKIL 119
Db 79 QIWTAGQERHTTTSYRGAMGIMLVYDITNGKSFENISKVLINIDEHANEDVERMLL 138
Qy 120 GNKADEKROVGRGQQLAKEYGMDIFYETSACTNLNIKESFTRLTVELVLAHREKLEGL 171
Db 139 GNKCDMDKRVVPKGGQIAREHGIRFFETSAKININIEKAFITLAEDILR 190
RESULT 13
ABB71647
ID ABB71647 standard; Protein; 207 AA.
XX ABB71647;
AC ABB71647;
XX 26-MAR-2002 (first entry)
DT Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
XX Drosophila melanogaster polypeptide SEQ ID NO 41733.
XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
PD 23-MAR-2001; 2001WO-US09231.
PF 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX N-PSDB; ABL15750.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 41733; 21pp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB27072).
XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 207 AA;
SQ Query Match 47.8%; Score 528.5; DB 22; Length 207;
Best Local Similarity 53.8%; Pred. No. 1.1e-47;
Matches 100; Conservative 41; Mismatches 40; Indels 5; Gaps 1;
Qy 1 MAKQYDVLFRLLIGSGVGKTCCLCRFTDNEPHSSHISTIGYDFKMKTIKIVGKIRIQ 60
Db 1 MAKTYDYLKLLIGSGVGKTCILFRFSDAFNTFTSTIGIDFKIRIETLKNKIKLQ 60
Qy 61 IWTAGQERYOTTTKYIRRAQGIIFLVYDITSSERSYQHIMKWYSDVDEVAPEGVKILIG 120
Db 61 IWTAGQERFTTTAYRGAMGIMLVYDITQEKSFENIKWIRNIEENASADVERKMLLG 120
Qy 121 NKADEKROVGRGQQLAKEYGMDIFYETSACTNLNIKESFTRLTVELVLAHREKLEGL 180
Db 121 NKGLTDRKQVSRGEQQLAIEYGIKFMETSASAKINVEEAF-----LTLSIDIKATEK 175
Qy 181 RMRASN 186
Db 176 RMEANN 181
RESULT 14
AAG67154
ID AAG67154 standard; Protein; 207 AA.
XX AAG67154;
AC AAG67154;
XX 13-NOV-2001 (first entry)
DT Amino acid sequence of human 27423 G-protein.
XX Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
XX congenital anomaly; pulmonary congestion; oedema; haemorrhage;
KW

Search completed: November 17, 2002, 20:38:26
Job time : 68 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2002, 20:39:14 ; Search time 2665 Seconds
(without alignments)
2315.120 Million cell updates/sec

Title: US-09-817-198a-2
Perfect score: 1105
Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEERKPEGPANSSKTCWC 212

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09817198/runat_13112002_135207_7033/app_query.fasta_1.391
-DB=GenEmbl -QFWT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DUALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09817198.ecgn.1.1616.qrunat_13112002_135207_7033 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THRAIDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.in:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg_hum:*
31: em.htg_inv:*
32: em.htg_other:*
33: em.htg_mus:*
34: em.htg_pln:*
35: em.htg_rod:*
36: em.htg_mam:*
37: em.htg_vrt:*
38: em.sy:*
39: em.htgo_hum:*
40: em.htgo_mus:*
41: em.htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1105	100.0	1054	6	AX399903 Sequence
2	1090	98.6	3139	10	BC027769 Mus muscu
3	1077	97.5	945	10	M83679 Sprague-Daw
4	798	72.2	2560	10	BC013790 Mus muscu
5	546.5	49.5	1613	3	AB006189 Drosophil
6	546.5	49.5	2349	3	AY060425 Drosophil
7	545	49.3	624	9	AF498943 Homo sapi
8	545	49.3	660	9	AF498943 Homo sapi
9	545	49.3	760	4	X56385 Canine rab8
10	545	49.3	2048	9	BC002977 Homo sapi
11	545	49.3	2818	9	AK025165 Homo sapi
12	541	49.0	1337	10	BC019990 Mus muscu
13	539.5	48.8	765	5	DYGORA2
14	534.5	48.4	1211	9	BC000799 Homo sapi
15	534.5	48.4	1238	9	HSRAB13
16	532	48.1	612	9	AF498948 Homo sapi
17	531.5	48.1	840	5	DYGORAL
18	530.5	48.0	603	9	AF297660 Homo sapi
19	530.5	48.0	603	9	AF498945 Homo sapi
20	530.5	48.0	888	4	CFRAB10
21	530.5	48.0	897	10	AF035646 Mus muscu
22	530.5	48.0	1029	9	BC000896 Homo sapi
23	530.5	48.0	3164	9	AF106681 Homo sapi
24	530.5	48.0	3533	9	AK023223 Homo sapi
25	529.5	47.9	954	8	ATHARA3
26	529.5	47.9	969	8	AY042795 Arabidopsi
27	529.5	47.9	1143	8	AY035132 Arabidops
28	528.5	47.8	624	6	AX236078 Sequence
29	528.5	47.8	740	10	RNU53475 Rattus norv
30	528.5	47.8	759	10	AF525280 Rattus no
31	528.5	47.8	1128	9	BC020654 Homo sapi
32	528.5	47.8	1161	6	AX236076 Sequence
33	528.5	47.8	1265	9	AB038995 Homo sapi
34	528.5	47.8	2048	3	AY069671 Drosophil
35	528.5	47.8	2210	3	D84347 Drosophila
36	528.5	47.8	2497	6	AX285074 Sequence
37	528.5	47.8	2497	6	AX285080 Sequence
38	528.5	47.8	2497	6	AX285089 Sequence
39	528.5	47.8	2497	6	AX285091 Sequence
40	528.5	47.8	3077	9	AK001111 Homo sapi
41	528.5	47.8	3110	9	HSM804678 Homo sapi
42	526.5	47.6	1113	8	LJRBABE
43	526.5	47.6	3521	9	HSM801620 Homo sapi
44	526	47.6	638	10	S53270 MEL-RAS-rel
45	526	47.6	1118	10	BC027214 Mus muscu

ALIGNMENTS


```
QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
Db 700 TCCTTCACTCGTCTGACGAGCTGGTCTGCAGGCTCACAGAAAGAGCTGGATGGTCTC 759
QY 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGlyLysPro 200
Db 760 CGAATATGTCACCAATAGCTCGCACTGGCCGAGCTGGAGGAAGACGAAAGCAAACT 819
QY 201 GluGlyProAlaAsnSerSerLysThrCysTrpCys 212
Db 820 GAAGGCCCCAGCAAACTCTTCAAGACCTGCTGGTGC 855

RESULT 4
BC013790 2560 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, clone MGC:6897 IMAGE:2655151, mRNA, complete cds.
DEFINITION BC013790
ACCESSION BC013790.1 GI:15489393
VERSION MGC.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 2560)
JOURNAL Strausberg,R.
Direct Submission
Submitted (04-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps@remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Yoonarntne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Xuan, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 5 Row: i Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction.
Location/Qualifiers
1. .2560
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N-3"
/clone="MGC:6897 IMAGE:2655151"
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
old mouse. Taken by biopsy."
/clone_lib="NCL CGAP_Mam2"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
134. .660
/codon_start=1
/product="Unknown (protein for MGC:6897)"
/protein_id="AAH13790.1"
/db_xref="GI:15489394"
/translation="MAKQIDVLFRLLLGDSGVKGTCLLCRTDNEFHSSHSITIGVD
FKMKTDVIGIKVRIQIWDTAGERYQITIKQYRRAQGIPLVYDISSERSYQHINKW
VSDVDEYAPGVQKILIGNKADEQKRVGREGQQLAKEYGMDFEYTSACPFSSFFS
FSTPSPL"

BASE COUNT 614 a 648 c 719 g 579 t
ORIGIN
```

Alignment Scores:

pred. No.: 1 65e-69 Length: 2560
Score: 798.00 Matches: 152
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.35% Mismatches: 0
Query Match: 72.22% Indels: 0
DB: 10 Gaps: 0

US-09-817-198a-2 (1-212) x BC013790 (1-2560)

QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
Db 154 ATGCGAAACAGTACGATGTCTGTCGGCTACTGTCGGGACTCCGGGTGGC 213
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db 214 AAGACATGCTGCTGTGCGCTTCCACGACACAGAGTCCACTCCTCGCATATCTCCACC 273
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 274 ATCGGTGTTGACTTTAAGATGAAGACTATCGATGTAGACGGCATCAAGTGAGANTACAG 333
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 334 ATTTGGGACACAGCAGCGGAGGAGGTACAGACTATCACAAAGCAGTACTATCGGCGA 393
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 394 GCCCAGGGAATATTTTAGTCTACGACATTAGCAGTGAGCGCTCCTATCAGCATATCATG 453
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 454 AACTGGGTGACGTGAGTGTAGTACGCTCCAGAGGAGTCCAGAGATCCTTAATTGGG 513
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
Db 514 AATAAGGCTGTGTAAGAGCAGAAACGCGAGTGGGGAGAGAGCGGCGCAGCAGCTGGCT 573
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCys 153
Db 574 AAGGATACGGCATGACCTTCTACGAACAAGTGCCTGC 612

RESULT 5

AB006189 1613 bp mRNA linear INV 07-AUG-1997
LOCUS Drosophila melanogaster mRNA for Rab10, complete cds.
DEFINITION AB006189
ACCESSION AB006189.1 GI:2317271
VERSION Rab10.
KEYWORDS Drosophila melanogaster (strain:Oregon R) head cDNA to mRNA.
SOURCE ORGANISM Drosophila melanogaster
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS 1 (sites)
TITLE Rab proteins of Drosophila melanogaster: novel members of the
Rab-protein family
JOURNAL FEBS Lett. 404 (1), 65-69 (1997)
MEDLINE 9728579
REFERENCE 2 (bases 1 to 1613)
AUTHORS Ozaki,K.
TITLE Direct Submission

JOURNAL

Submitted (01-AUG-1997) Koichi Ozaki, Osaka University, Graduate
School of Science, Department of Biology, 1-1 Machikaneyama,
Toyonaka, Osaka 560, Japan (E-mail:ozaki@bio.sci.osaka-u.ac.jp,
Tel:+81-6-850-5439, Fax:+81-6-850-5439)

FEATURES

Location/Qualifiers
1. .1613
/organism="Drosophila melanogaster"
/strain="Oregon R"
/db_xref="taxon:7227"

RESULT 6
AY060425
LOCUS

REFERENCE AUTHORS TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea. 1 (bases 1 to 660) Zahraoui, A. Direct Submission Submitted (26-NOV-1990) A. Zahraoui, INSERM-U 248, 10 AVENUE DE VERDUN, 750-10 PARIS, FRANCE
REFERENCE AUTHORS TITLE JOURNAL	2 (bases 1 to 660) Zahraoui, A., Joberty, G., Arpin, M., Fontaine, J. J., Hellio, R., Tavtavian, A. and Louvard, D. A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cells but colocalizes with the tight junction marker ZO-1 in polarized epithelial cells J. Cell Biol. 124 (1-2), 101-115 (1994)

FEATURES	source	Location/Qualifiers
		1..660
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone_lib="Human pheochromocytoma cDNA library"
gene		11..634
		/gene="rab8"
CDS		11..634
		/gene="rab8"
		/codon_start=1
		/product="rab8 small GTP binding protein"
		/protein_id="CAA0065.1"
		/db_xref="GI:452318"
		/db_xref="SWISS-PROT:P24407"
		/translation="MAKTYDYLFKLLIGDSGVGKTCVLFRFSEDAFNSTFTSTIGID FKRIELGGRIKLIKQWTAGCERFTITAYIRGAMGLMVLVDITNKSFFNIRNW IKNIIEHASADVEKMLIGNDVNDKRVSKERGEKALDYGKFMETSAKANINVEN AFPTFLIARDITKAKMDKKLEGNPQSGNQSKVITPDQQRSSFFRCVLL"
BASE COUNT	197 a	159 c 172 g 132 t

Alignment Scores:		
Pred. No.:	3,64e-45	660
Score:	545.00	99
Percent Similarity:	77.01%	Matches: 45
Best local Similarity:	52.94%	Mismatches: 43
Query Match:	49.32%	Indels: 0
DB:	9	Gaps: 0

US-09-817-198A-2 (1-212) x HSMRAB8 (1-660)

1	MetalAlaysGlnTryAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly	20
QY		
11	ATGGCGAAGACCTACGATTACCCTTTCAAGTCGTCTGATCGGGACTCGGGGTGGG	70
Db		
21	LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr	40
QY		
71	AAGACCTGTGTCTTGTTTCCGCTTCTCCGAGAGCGCTTCACTCCACTTTTATCTCCAC	130
Db		

QY	41	lleGlyValaspPheIysMetIysThrIleGluValaspGlyIleLysValArgIleGln	60
Db	131	ATAGCAATTGACTTTAAATTAGGACCATAGAGCTCGATGTCAGAAATTTAACTGCAG	190
QY	61	lletraspthrAlaGlyIleGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg	80

81	AlaGlnGlyIlePheLeuVal	ValTrpAspIleSerSerGluArgSer	GlnHisIleMet	1000
Qy				
251	GCAATGGCAATCAATGGTCTACAGACATACCAACAGAGAAGTCCTTCGACACATCCGG			310
Db				

[illegible]

Qy	141	LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu	160
		::: ::: ::::::	
Db	431	CTCGACTATGGAATCAAGTTTCATGGAGCACCAGCGCAACATCAATGTGGAAT	490
Qy	161	SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu	180
		::: ::::	
Db	491	GCATTTTTCATCTTCGCCAGAGATATCAAGACGAAAAATGGACAAAAAATTTGGAAGGCAAC	550
Qy	181	ArgMetArgAlaSerAsnGlu	187
		::: :::	
Db	551	AGCCCCAGGGGAGCAACCAAG	571
RESULT 9			
CFRAB8			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
SUBMITTED (17-AUG-1990)			
Zerial M., EMBL, Meyerhofstrasse 1, 6900			

```

FEATURES
source
  Heidelberg, FRG
  Location/Qualifiers
    1..760
      /organism="Canis familiaris"
      /strain="cocker spaniel"
      /db_xref="taxon:9615"
      /clone="2"
      /cell_line="MDCK"

```

```

gene
10..633
/genes="rab8"
CDS
10..633
/genes="rab8"
/function="GTP binding protein"
/codon_start=1
/protein_id="CAB56776.1"
/db_xref="GI:6006436"
/translation="MAKTYDYLKLLIGDSVGKTCVLPFRSEDAFNSTFTIGID
FKRIELDGKRIKLQWDTAQGERFTITAYRGAMGIMLYDITNKSFDNIRNW
IRNIEIHAASADVEKMTILGNKCDVNDKRRQSKERGEKALDYGKFMETSAKANINVEN
AFPTLTARDLKAKMDKKLEGNSQGSNQGVKITPDQQRKRSFFRCVLL"
215 a 190 c 196 g 159 t
BASE COUNT
ORIGIN

```

Alignment Scores:	
Pred. No.:	4,32e-45
Score:	545.00
Length:	760
Matches:	99
Conservative:	45
Mismatches:	43
Indels:	0
Gaps:	0
DB:	4

US-09-817-198A-2 (1-212) x CFRB8 (1-760)

Oy	1	MetalysGlnTrAspValLeuPheArgLeuLeuLeuLeuLeuValcyl	20
Db	10	ATGGCGAAGACCTACGATTACCTGTTCACACTGCTGATCGGGAGCTCGGGGGTGGGG	69
Ov	21	LvsThrCvsLeuLeuCysAcgPheThrAspAsnGluPheHisSerSerHisIleSerThr	40

```
Db 70 AAGACTGTGCTGCTGCTCCGAGGAGCGCTTCAACATCAACTTTCATCTCCACT 129
QY 41 ILeGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIleGln 60
Db 130 ATAGGAATGACTTTAAATAGACCATAGACTCGATGGCAAGAGAAATAGCTACAG 189
QY 61 ILeTrrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 190 ATATGGGACACAGCTGGTCAAGAACGGTTTCGGACGATCACACACAGCCATTACAGGGC 249
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 250 GCAATGGGCATCATGCTGTCTATGACATCACCAACGAGAAATCCCTTGACAATATCCGG 309
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 310 AACTGGATTCGGACATTCAGGAGCATGCTTCGACAGATGTCGAAAGATGATCTCGGA 369
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnLeuAla 140
Db 370 AACAACTGTGATGTGAACGACAAAGACAAAGTTTCCAAAGGAGCGGAGAAAGCTGGCC 429
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 430 CTGCACTATGGAATCAAGTTTCATGAGACCATGTCGAAAGCCCAACATCAATGTCGAGA 489
QY 161 SerPheThrArgLeuThrGluLeuValGlnAlaHisArgLysGluLeuGluGlyLeu 180
Db 490 GCATTTTCTACTCTCCGAGAGACATCAAGCAAGCAAGATGGACAAAATTTGGAGGCAAC 549
QY 181 ArgMetArgAlaSerAsnGlu 187
Db 550 AGTCCCAAGGAGCAACACAG 570

RESULT 10
BC002977 2048 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, mel transforming oncogene (derived from cell line
DEFINITION NK14)- RAB8 homolog, clone MGC:2196 IMAGE:3547214, mRNA, complete
cids.
ACCESSION BC002977
VERSION BC002977.1 GI:12804236
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2048)
Strausberg,R.
Direct Submission
Submitted (06-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhui, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasia van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
```

```
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 8 Row: d Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 452317.
FEATURES
    source
        Location/Qualifiers
            1..2048
                /organism="Homo sapiens"
                /db_xref="LocusID:4218"
                /db_xref="taxon:9606"
                /clone="MGC:2196 IMAGE:3547214"
                /tissue_type="Skin, melanotic melanoma."
                /clone_lib="NIH_MGC_20"
                /lab_host="DH10B-R"
                /note="Vector: pOTB7"
            77..700
                /codon_start=1
                /product="mel transforming oncogene (derived from cell
                line NK14)- RAB8 homolog"
                /protein_id="AAH02977.1"
                /db_xref="GI:12804237"
                /translation="MAKTYDYLFKLLLLIGDSGVGKTVLRFSEDAFNSTFTSTIGID
                FKIRTELDGKRIKLIQINDTAGOERFTTAYYRGAMGIMLVYDITNKSNDINRW
                INKIEHASADVEKMLGNKQVNDKRVSKERGEKALDYGIKFMETSAKANINVEN
                AFPTLARDIKAKMDKKLGNKSPQSNQGVKITPDQOKRSSFFRCVLL"
BASE COUNT 532 a 556 c 493 g 467 t
ORIGIN
Alignment Scores:
Pred. No.: 1,45e-44 Length: 2048
Score: 545.00 Matches: 99
Percent Similarity: 77.01% Conservative: 45
Best Local Similarity: 52.94% Mismatches: 43
Query Match: 49.32% Indels: 0
DB: 9 Gaps: 0
US-09-817-198a-2 (1-212) x BC002977 (1-2048)
QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
Db 77 ATGCGAGACCTTACCATTTACCTGTTCAAGCTGCTGATCGGGGACTCGGGGGTGGG 136
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
Db 137 AAGACCTGTCTCTGCTTCTCGCTTCTCGAGGAGCGCTTCAACTCCACTTTATCTCCACC 196
QY 41 ILeGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIleGln 60
Db 197 ATAGGAATGACTTTAAATAGACCATAGACTCGATGGCAAGAGAAATTAAGTCCAG 256
QY 61 ILeTrrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 257 ATATGGGACACAGCGGTCAGGAACGGTTTCGGACGATCACACAGCGCTTACTACAGGGT 316
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 317 GCAATGGGCATCATGCTGTCTACGACATCACCAACGAGAGAGTCTTCGACAAACATCCG 376
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 377 AACTGGATTCGCAACATTTAGGAGACGACGCTCTGCGACACGTCGAAAAGATGATCTCGG 436
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnLeuAla 140
Db 437 AACAACTGTGATGTGATGACAGACAGACAAAGTTTCCAAAGGACGCGGAGAAAGCTGCC 496
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 497 CTCGACTATGGAATCAAGTTTCATGAGACGACGCGGAGGCAACATCAATGTCGAAAAT 556
QY 161 SerPheThrArgLeuThrGluLeuValGlnAlaHisArgLysGluLeuGluGlyLeu 180
Db 557 GCATTTTCTACTCTCGCCAGAGATATCAAGCAAGAAAATTTGGAGGCAAC 616
```


[illegible]


```

QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysLeuLeuGly 120
      |||:::  :::::|||||::|
Db 314 AACTGGATAAGGAATATTCAAGAGCATGCTTCAGATGTAGAAAGATGATTCTGGGC 373

QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
      |||::|  |||  ::|||::|::|::|::|::|::|::|::|::|::|::|
Db 374 ACAAGTGGCAGCATGAATGAAAAAGGCAAGTGTCAAAGCAACGAGGAGAAAGTAGCG 433

QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnLysGlu 160
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 434 ATAGATTATGGATAAAGTCTCTGAAACCCAGCGCAAAATCCAGCATAAACGTTGAAGAG 493

QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
      ::|||  |||  :::::
Db 494 GCATTATTACACTTGCACGAGATATCATGACGAACTCAACAAGAAATG----- 544

QY 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGlu---LeuGluGluGluGlyLys 199
      |||::|  ::|||  ::|||  |||  :::::
Db 545 -----AATGAAACAGTCTCCAGGAGGCGAGTGAACAAAGTTAAATCACCG 589

QY 200 ProGluGlyProAlaAsnSerSerLys 208
      |||:::  |||:::
Db 590 CCAAGAACCCAGTCAAAAGAAAG 616

RESULT 14
LOCUS BC000799 1211 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, RAB13, member RAS oncogene family, clone MGC:5074
IMAGE:3451945, mRNA, complete cds.
ACCESSION BC000799
VERSION BC000799.1 GI:12654002
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1211)
Strausberg,R.
Direct Submission
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxii.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAP Plate: 4 Row: 1 Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4506362.
Location/Qualifiers
1. 1211
/organism="Homo sapiens"
/db_xref="LocusID:5872"
/db_xref="taxon:9606"
/clone="MGC:5074 IMAGE:3451945"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NTH_MGC_10"
/lab_host="DH10B"
/Note="Vector: pCMV-SPORT6"
95. .706

FEATURES
source
CDS

```

```

/codon_start=1
/product="RAB13, member RAS oncogene family"
/protein_id="AAH00799.1"
/db_xref="GI:12654003"
/translation="MAKAYDHLFKLLLLDGSVGVKTCILIRFAEDFNNTYISTIGID
FKIRTVDIGKKIKLVQWDTAGREFKTTTAYIRGAMGIIILYDITDEKSENIQNW
KSKTENASAGVERLLNKCMEARKVQEQADKLAREHGIRFEFFETSAKSSMNVDE
AFSSLIARDILLKSGRRSGNGKNKPPSTDLKTCDKKNTKCSLG"
BASE COUNT 392 a 256 c 335 g 228 t
ORIGIN
Alignment Scores:
Pred. No.: 8.38e-44 Length: 1211
Score: 534.50 Matches: 102
Percent Similarity: 66.38% Conservative: 54
Best Local Similarity: 43.40% Mismatches: 55
Query Match: 48.37% Indels: 25
DB: 9 Gaps: 1
US-09-817-198a-2 (1-212) x BC000799 (1-1211)
QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly 20
      |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 95 ATGGCCAAAGCCCTACGACCACCTCTTCAAGTTGCTGCTGATCGGGGACTCGGGGGTGGC 154

QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisSerThr 40
      |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 155 AGACTGTGTCATCATTCGCTTCGAGAGGACAACTTCAACAACACTTACATCTCCACC 214

QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgLysGln 60
      |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 215 ATCGGAATTTGATTTCAAGATCGGCACTGTGATATAGAGGAGAAAGATCAAACTACAA 274

QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTrpArg 80
      |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|  |||::|
Db 275 GTCTGGGACACGGCTGGCCAGAGCGGTTCAGACAACTAATCACTACCTACCGGTGGA 334

QY 81 AlaGlnGlyIlePheLeuValTyrAspLysSerSerGluArgSerTyrGlnHisIleMet 100
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 335 GCATCGGGCATTTATCTTAGTATACGACATCAGGATGAGAAATCTTTCAGAAATATTCAG 394

QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysLeuLeuGly 120
      |||:::  ::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 395 AACTGGATGAAAGCATCAAGGAGATGCTCGGCTGGGTGGGCGCTCTTGCTGGGG 454

QY 121 AsnLysAlaAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
      |||::|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 455 AACAAATGTGACATGGAGGCCCAAGAGAGGTGCGAAGAGGACGAGCCGATAGTTGGCT 514

QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnLysGlu 160
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 515 CGAGAGCATGGAATCGGATTTTCGNAACTAGTCTAAATCCAGTATGATGTGGATGAG 574

QY 161 SerPheThrArgLeuThrGluLeuValLeu----- 170
      ::|||:::  |||  ::|||
Db 575 GCTTTTAGTTCCTGCGCCGGGACATCTTGTCTCAAGTCAGAGCGCGGAGATCAGGAAAC 634

QY 171 -----GlnAlaHisArgLysGlu 176
      ::|  |||:::
Db 635 GGCACAAAGCCCTCCAGTACTGACCTGAAAACCTGTGACAAAGAAAG-CACCAACAAAGTG 693

QY 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 694 CTCCTCGGCTGAGGACCCCTTCTTCCTCCACCACCCCGGAGCTGAACCTGAGGGAGAC 753

QY 197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211
      |||:::  ::|||:::
Db 754 AACGGCAGGAGGTGAGCAGGAGGAGAAATAGCAGAGGGGGCTTGG 798

RESULT 15
HSRAB13
LOCUS HSRAB13
DEFINITION H.sapiens mRNA for rab 13.
1238 bp mRNA linear PRI 02-FEB-1994

```

Db	440	AAC	TGGATGAAAGCATCAAGGAGATCCCTCGCTGGGTGAGCGCCTCTTGCTGGGG	499
Qy	121	Asn	LysAlaaspGluGluGlnLysArgGlnValGlyArgGluGlnGlnLeuAla	140
Db	500	AACAAATGTGACATGGAGCCAGAGGAAGGTGCAGAAAGCAGCGCATGAATGGCT	559	
Qy	141	Lys	GluTyrGlyMetaspPheTyrGluThrSerAlaCysThrAsnLeuasnLeLysGlu	160
Db	560	CGAGAGCATGGAATCCGATTTTCGAAACTAGTGTCTAAATCCAGTATGAATGTGGATGAG	619	
Qy	161	Ser	PheThrArgLeuThrGluLeuValLeu-----	170
Db	620	GCTTTTAGTTCCCTGGCCCGGAGCATCTTGCTCAAGTCAGGAGCGCGAGATCAGGAAC	679	
Qy	171	-----	-----GlnAlaHisArgLysGlu	176
Db	580	GGCAACAAAGCCTCCAGTACTGACGTGMAAACTGTGACAGAAGAA-CACCAACAAGTG	738	
Qy	177	Leu	GluGlyLeuArgMetArgAsnGluLeuAlaLeuAlaGluGluGlu	196
Db	739	CTCCCTGGGTGAGGACCTTTCTTGCCTCCCAACCCGGAAGCTGAACCTGAGGGAGAC	798	
Qy	197	Glu	GlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp	211
Db	799	AACGGCAGGGAGTGAGCAGGGGAGAAATAGCAGAGGGGCTTGG	843	

Search completed: November 17, 2002, 21:32:05
Job time : 2670 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 18:07:04 ; Search time 67 Seconds
(without alignments)
421.629 Million cell updates/sec

Title: US-09-817-198A-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLLLIGDSGVG.....LEEEGKPGPANSSKTCWC 212

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1105	100.0	401	AAU17136	Novel signal trans
2	1092	98.8	218	AAU17136	Human ORF1368
3	832	75.3	188	AAU17555	Novel signal trans
4	546.5	49.5	204	AB070670	Drosophila melanog
5	545	49.3	213	AB070670	Human rab8 homolog
6	540	48.9	221	ABP41333	Human ovarian anti
7	532	48.1	246	AA198196	Lung cancer associ
8	530.5	48.0	200	AA19165	Amino acid sequenc
9	530.5	48.0	200	AA099979	Human Rab10 protei
10	530.5	48.0	200	AA095340	Human protein sequ

11	530.5	48.0	218	21	AA056993	Human prostate can
12	529.5	47.9	218	22	AB07266	Novel human diagno
13	528.5	47.8	207	22	AB071647	Drosophila melanog
14	528.5	47.8	207	22	AB071647	Amino acid sequenc
15	528.5	47.8	207	22	AB071647	Human protein sequ
16	526	47.6	201	21	AA09982	Canine Rab10 prote
17	523	47.3	199	21	AA09980	Human Rab10 protei
18	523	47.3	199	21	AA09981	Human Rab10 protei
19	519	47.0	216	21	AA08688	Arabidopsis thalia
20	519	47.0	216	21	AA08688	Arabidopsis thalia
21	519	47.0	253	21	AA053944	Arabidopsis thalia
22	519	47.0	254	21	AA08687	Arabidopsis thalia
23	513.5	46.5	215	21	AA05215	Zea mays protein f
24	510	46.2	218	21	AA019220	Arabidopsis thalia
25	507	45.9	216	21	AA08006	Arabidopsis thalia
26	506.5	45.8	209	22	AB023365	Novel human diagno
27	501	45.3	224	21	AA047826	Arabidopsis thalia
28	501	45.3	234	21	AA047825	Arabidopsis thalia
29	501	45.3	335	22	AB023366	Novel human diagno
30	478.5	43.3	201	20	AA000919	Human Rab protein,
31	478.5	43.3	201	22	AA028024	Novel human secret
32	478.5	43.3	221	22	AA013525	Human polypeptide
33	478.5	43.3	224	21	AA058758	Breast and ovarian
34	476	43.1	203	21	AA030499	Arabidopsis thalia
35	476	43.1	258	21	AA030498	Arabidopsis thalia
36	468	42.4	202	21	AA007763	Gene 44 human secr
37	467	42.3	205	21	AA034843	Human secreted pro
38	467	42.3	205	21	AA034844	Arabidopsis thalia
39	466	42.2	202	21	AA010858	Arabidopsis thalia
40	462	41.8	205	22	AB059808	Drosophila melanog
41	457	41.4	218	21	AA030710	Arabidopsis thalia
42	457	41.4	221	21	AA030709	Arabidopsis thalia
43	457	41.4	254	21	AA053523	Arabidopsis thalia
44	456.5	41.3	227	21	AA019163	Amino acid sequenc
45	456.5	41.3	227	22	AA039600	Human polypeptide

ALIGNMENTS

RESULT 1

AAU17136
ID AAU17136 standard; Protein; 401 AA.

AC AAU17136;

DT 07-NOV-2001 (first entry)

DE Novel signal transduction pathway protein, Seq ID 701.

DE Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.

OS Homo sapiens.

PN WO200154733-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01312.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225477.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234597.
 PR 25-SEP-2000; 2000US-0234598.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244517.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249266.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI: 2001-465460/50.
 DR N-PSDB; AAS27053.
 XX
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders
 XX
 PS Claim 1; SEQ ID No 701; 880pp; English.
 XX

CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX

Query Match 100.0%; Score 1105; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.1e-108;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKIVRIQ 60
DB 27 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKIVRIQ 86
QY 61 IWDTAGQERYQTITKQYRRAGQIFLVYDISSERSYQHIMKWVSDVDEYAPGVQKILIG 120
DB 87 IWDTAGQERYQTITKQYRRAGQIFLVYDISSERSYQHIMKWVSDVDEYAPGVQKILIG 146
QY 121 NKADEQRQVREGQOQOLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHRLKLEGL 180
DB 147 NKADEQRQVREGQOQOLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHRLKLEGL 206
QY 181 RMRASNELALAELEEEKGPEGPANSSKTCWC 212
DB 207 RMRASNELALAELEEEKGPEGPANSSKTCWC 238

RESULT 2
AAB41604
ID AAB41604 standard; Protein: 218 AA.

XX AC AAB41604;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF1368 polypeptide sequence SEQ ID NO:2736.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antipariatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

XX OS Homo sapiens.

XX PN WO200058473-A2.

XX 05-OCT-2000.
XX 31-MAR-2000; 2000WO-US08621.
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach M;
PI WPI: 2000-602362/57.
XX N-PSDB; RAC75813.
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX Claim 11; Page 1979-1980; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipariatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 218 AA;

Query Match 98.8%; Score 1092; DB 21; Length 218;
Best Local Similarity 97.2%; Pred. No. 2e-107;
Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKIVRIQ 60
DB 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKIVRIQ 60
QY 61 IWDTAGQERYQTITKQYRRAGQIFLVYDISSERSYQHIMKWVSDVDEYAPGVQKILIG 120
DB 61 IWDTAGQERYQTITKQYRRAGQIFLVYDISSERSYQHIMKWVSDVDEYAPGVQKILIG 120
QY 121 NKADEQRQVREGQOQO-----QLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHRL 174
DB 121 NKADEQRQVREGQOQOCPSPLOLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHRL 180
QY 175 KELEGLRMRASNELALAELEEEKGPEGPANSSKTCWC 212
DB 181 KELEGLRMRASNELALAELEEEKGPEGPANSSKTCWC 218

RESULT 3
AAU17555
ID AAU17555 standard; Protein: 188 AA.

XX AC AAU17555;

PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251989.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465460/50.
DR N-PSDB; AAS27472.
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
XX
PS Claim 1; SEQ ID NO 1120; 880pp; English.
XX
XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX

Query Match 75.3%; Score 832; DB 22; Length 188;
Best Local Similarity 97.6%; Pred. NO. 6.2e-80;
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 MAKQYDVLFRLLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDKFKMTIEVDGKVRIO 60
DB 24 MAKKYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDKFKMTIEVDGKVRIO 83
QY 61 IWDTAGQERYQITTKQYVYRAAGIFLVYDYSERSYQHIMKWVSDVDEYAPGVQKILIG 120
DB 84 IWDTAGQERYQITTKQYVYRAAGIFLVYDYSERSYQHIMKWVSDVDEYAPGVQKILIG 143
QY 121 NKADEQKRVQGVQGOQLAKEYGDMFYETSACTNLNKESTFR 164
DB 144 NKADEQKRVQGVQGOQLAKEYGDMFYETSACTNLNKESTFR 187

RESULT 4
ABB70670
ID ABB70670 standard; Protein; 204 AA.
XX

AC ABB70670;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 38802.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL14773.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 38802; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 204 AA;
Query Match 49.5%; Score 546.5; DB 22; Length 204;
Best Local Similarity 49.8%; Pred. NO. 1.4e-49;
Matches 105; Conservative 47; Mismatches 50; Indels 9; Gaps 3;
QY 1 MAKQYDVLFRLLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDKFKMTIEVDGKVRIO 59
DB 1 MAKKYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDKFKMTIEVDGKVRIO 60
QY 60 QIWDTAGQERYQITTKQYVYRAAGIFLVYDYSERSYQHIMKWVSDVDEYAPGVQKILIG 119
DB 61 QIWDTAGQERYQITTKQYVYRAAGIFLVYDYSERSYQHIMKWVSDVDEYAPGVQKILIG 120
QY 120 GNKADEQKRVQGVQGOQLAKEYGDMFYETSACTNLNKESTFRTELVLVLAHKELEG 179
DB 121 GNKCDMTKRVYVNGERGAIAIEHGRFMTSAKSNINIERAFCELAELAIID- ----KTSG 176
QY 180 LMRASNELALAELEEECKPEGPANSSKTC 210
DB 177 RESAENQERVIIDRRNQEKAP- ----GYSKCC 203

RESULT 5
ABB11916
ID ABB11916 standard; peptide; 213 AA.
XX
XX ABB11916;
AC

XX 11-JAN-2002 (first entry)
DE Human rab8 homologue, SEQ ID NO:2286.
KW Human: cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antisthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.
XX Homo sapiens.
OS WO200157188-A2.
PN 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US03800.
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
PI WPI: 2001-457740/49.
XX N-PSDB; ABA09160.
DR Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX Claim 20; Page 276; 1963pp; English.
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08223-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to

CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX Sequence 213 AA;
SQ Query Match 49.3%; Score 545; DB 22; Length 213;
Best Local Similarity 52.9%; Pred. No. 2.1e-49;
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;
QY 1 MAKQYDLFRLLIGDSGVGKTCCLCRFTDNEFHSIHSTIGVDFKMKTIYDVGKIVRIQ 60
DB III III III III III III III III III III III III III III III III
7 MAKTYDLFKLLIGDSGVGKTCVLFRESDFNSFTISIGIDFKIRTITELDGKRIKLQ 66
QY 61 IWDTAGQERYQTITKQYRRAGQIFLYVIDISSERSQHIKWKVSDVDYAPGKQKILIG 120
DB III III III III III III III III III III III III III III III III
67 IWDTAGQERFTTITAYYRGAMIMLVYDITNEKSFNIRNWIIRNIEHSAADVEKMIIG 126
QY 121 NKADEQKQVGRGEOGQOLAKKEYGDMFYETSACTNINLNIKESFTRLTLVLQAHKKEGL 180
DB II I IIII :
127 NKCDVNDKQVSKERGEKLDYGIKFMETSAKININVENAFFTLARDIKAKMDKKLEGN 186
QY 181 RMRASNE 187
DB : II :
187 SPOGSNQ 193
RESULT 6
ID ABB41333
XX ID ABB41333 standard; Protein; 221 AA.
AC ABB41333;
XX 23-AUG-2002 (first entry)
DT Human ovarian antigen HCGMA67, SEQ ID NO:2465.
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX Homo sapiens.
OS WO200200677-A1.
PN 03-JAN-2002.
XX 07-JUN-2001; 2001WO-US18569.
PR 07-JUN-2000; 2000US-209467P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Birse CE, Rosen CA;
PI WPI: 2002-147878/19.
DR N-PSDB; ABQ54410.
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -

XX PS Claim 11; SEQ ID No 2465; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The reference data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 221 AA;

Query Match 48.9%; Score 540; DB 23; Length 221;
Best Local Similarity 52.7%; Pred. No. 7.5e-49;
Matches 98; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Qy 2 AKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDPKMTIEVDGKIVRIQI 61
Db 16 AKTYDYLFRLLIGDSGVGKTCVLFVSEDAFNFTISTIGIDFKIRTVDEGKKIKLIQI 75

Qy 62 WDTAGQERYQTITTKQYRRAGQIFLYVDISSERSYQHIMKMWSDVDEYAPGVQKILIGN 121
Db 76 WDTAGQERFRTITTAAYRGAMGIMLVYDITNEKSPDNIRNIRNEEHSADVEKMILGN 135

Qy 122 KADEQKQVREGQOQLAKKEYGMDFYETSACTNLNKESTRTELVLQAHKLEGLR 181
Db 136 KCDVNDKQVSKERGEKIALDYGIKFMETSAKANINVENAFPTLARDIKAKMDKKLEGN 195

Qy 182 MRASNE 187
Db 196 PQGSNQ 201

RESULT 7
AAB58196
ID AAB58196 standard; Protein; 246 AA.

XX AC AAB58196;
XX DT 14-MAR-2001 (first entry)
XX DE Lung cancer associated polypeptide sequence SEQ ID 534.
XX KW Human; lung cancer associated protein; neuroprotective; cytostatic;
KW KW cardioactive; immunomodulatory; muscular active; vulnerary;
KW KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX KW proliferative disorder; wound healing; infectious disease.
XX OS Homo sapiens.

XX WO200055180-A2.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05918.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX Ruben SM;
XX WPI: 2000-587514/55.
XX N-PSDB: AAF18072.
XX Lung cancer associated gene sequences, referred to as lung cancer
XX PT antigens, useful for treatment, prevention, and diagnosis of disorders
XX PT such as lung cancer -
XX PS Claim 11; Page 1027-1028; 1425pp; English.
XX CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.

XX SQ Sequence 246 AA;

Query Match 48.1%; Score 532; DB 21; Length 246;
Best Local Similarity 47.6%; Pred. No. 6.2e-48;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDPKMTIEVDGKIVRIQ 60
Db 44 MAKAYDHLFKLLIGDSGVGKTCCLIRFADNFNNTYISTIGIDFKIRTVDEGKKIKLIQ 103

Qy 61 IWDTAGQERYQTITTKQYRRAGQIFLYVDISSERSYQHIMKMWSDVDEYAPGVQKILIG 120
Db 104 VMDTAGQERFRTITTAAYRGAMGIMLVYDITNEKSPDNIRNIRNEEHSADVEKMILGN 163

Qy 121 NKADDEQKQVREGQOQLAKKEYGMDFYETSACTNLNKESTRTELVLQAHKLEGLR 180
Db 164 NKCDMEAKRKVQKEQADKLAREHGIRFFETSAKSMNVDEAFSSLDAIL-----LASKG 217

Qy 181 RMRASNELALAELEEEEGKPEGPANSSKTC 210
Db 218 GRRSGN-----GNKP--PSTDCLKTC 235

RESULT 8
AAB19165
ID AAB19165 standard; Protein; 200 AA.
XX AC AAB19165;
XX DT 19-FEB-2001 (first entry)
XX

```
DE Amino acid sequence of human RAB10.
KW RAB protein; GTPase; GTP binding; gene therapy; cancer.
OS Homo sapiens.
XX
PN WO200058464-A2.
XX
PD 05-OCT-2000.
XX
PF 13-MAR-2000; 2000WO-US06330.
XX
PR 25-MAR-1999; 99US-0126083.
XX
PA (AXYS-) AXYS PHARM INC.
PI Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;
XX
DR WPI: 2000-647233/62.
DR N-PSDB; AAA96887.
XX
XX Novel isolated nucleic acid encoding a mammalian RAB protein useful for
PT identifying homologous or related genes, in producing composition that
PT modulates expression or function of RAB for cancer therapy -
XX
PS Claim 2; Page 39; 58pp; English.
XX
CC The present sequence represents a mammalian RAB protein. RAB proteins
CC constitute the largest family of small GTPases, with over 40 identified
CC isoforms. RAB proteins contain four highly conserved peptide sequences
CC involved in GTP binding and hydrolysis. Compositions comprising RAB
CC nucleic acid are useful for identifying homologous or related genes,
CC in producing compositions that modulate the expression or function of
CC RAB, for gene therapy, mapping functional regions of the protein and
CC in studying associated physiological pathways. In addition, modulation
CC of the gene activity in vivo is used for prophylactic and therapeutic
CC purposes, such as treatment of cancer, and identification of cell type
CC based on expression. The DNA may also be used to identify expression of
CC the gene in a biological specimen.
XX
XX Sequence 200 AA;
XX
Query Match 48.0%; Score 530.5; DB 21; Length 200;
Best Local Similarity 55.8%; Pred No. 6.6e-48;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKQ-YDVLFRLLIGDVGKTCCLCRFTDNEFHSHSTIGVDPKMTIEVDGIKVI 59
DB 1 MAKTTYDLLFKLLIGDVGKTCVLFRESDDAFNTFTSTIGIDFKIKTVELQKKIKL 60
QY 60 QIWDTAGQERYQTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILI 119
DB 61 QIWDTAGQERFHTTTSYRGANGIMLVYDITNGKSFENISKWLNRNIDSHANEDVERMLL 120
QY 120 GNKADEQKRVQREGQQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQ 171
DB 121 GNKCDMDKRRVVPKGEQIAREHGIRFETSAKANINIEKAFLTLAEDILR 172
RESULT 9
AAB09979
ID AAB09979 standard; Protein; 200 AA.
XX
AC AAB09979;
XX
DT 19-OCT-2000 (first entry)
DE Human Rab10 protein.
XX
KW Human; Rab10; Yap/Rab family.
XX
OS Homo sapiens.
XX
```

```
PN CN1249345-A.
XX
PD 05-APR-2000.
XX
PF 28-SEP-1998; 98CN-0121911.
XX
PR 28-SEP-1998; 98CN-0121911.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Tu Q, Gao J;
XX
DR WPI: 2000-400723/35.
DR N-PSDB; AAA40104.
XX
XX Preparation of human gene coding sequence, its encoded polypeptide -
PT Claim 2; Page 18; 23pp; Chinese.
XX
CC This invention describes a novel human Rab10 cDNA sequence. The protein
CC coded by said sequence belongs to Yap/Rab family and is the homolog of
CC mouse Rab10. The present invention also relates to the polypeptide coded
CC by said nucleotide sequence and the application and preparing process
CC of said polynucleotide and said polypeptide. This sequence represents the
CC human Rab10 protein described in the method of the invention.
XX
XX Sequence 200 AA;
XX
Query Match 48.0%; Score 530.5; DB 21; Length 200;
Best Local Similarity 55.8%; Pred No. 6.6e-48;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKQ-YDVLFRLLIGDVGKTCCLCRFTDNEFHSHSTIGVDPKMTIEVDGIKVI 59
DB 1 MAKTTYDLLFKLLIGDVGKTCVLFRESDDAFNTFTSTIGIDFKIKTVELQKKIKL 60
QY 60 QIWDTAGQERYQTITKQYRRAGQIFLVYDISERSYQHIMKWSVDVEYAPGVQKILI 119
DB 61 QIWDTAGQERFHTTTSYRGANGIMLVYDITNGKSFENISKWLNRNIDSHANEDVERMLL 120
QY 120 GNKADEQKRVQREGQQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQ 171
DB 121 GNKCDMDKRRVVPKGEQIAREHGIRFETSAKANINIEKAFLTLAEDILR 172
RESULT 10
AAB95340
ID AAB95340 standard; Protein; 200 AA.
XX
AC AAB95340;
XX
DT 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:17619.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
```


PD 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
DR N-PSDB; AAS71453.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID NO 37625; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 218 AA;
Query Match 47.9%; Score 529.5; DB 22; Length 218;
Best Local Similarity 55.8%; Pred. No. 9.6e-48;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKO-YDVLFRLLIGDSGVGKTCILCRFTDNEPHSSHISTIGVDFKMKTIYVDGKIVRI 59
DB 19 MAKRTYDLFLKLLIGDSGVGKTCILCRFTDNEPHSSHISTIGVDFKMKTIYVDGKIVRI 78
QY 60 QIWDTAGOERYQTITTKOYIRRAOGIFLVYDISSERSYQHIMKWVSDYDEVAPEGVQKILI 119
DB 79 QIWDTAGOERFHTITTYIRGANGIMLVYDITNGKSFENISKVLINIDEHANEVDVERMLL 138
QY 120 GNKADEQKROGVREGOQOLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHKLEGL 171
DB 139 GNKCDMDKRVVPKGKEQIAREHGIRFFETSAKANINIEKAFLTLAEDILR 190
RESULT 13
ID ABB71647
XX ABB71647 standard; Protein; 207 AA.
XX ABB71647;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 41733.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW

KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL15750.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 41733; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABJ01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 207 AA;
Query Match 47.8%; Score 528.5; DB 22; Length 207;
Best Local Similarity 53.8%; Pred. No. 1.1e-47;
Matches 100; Conservative 41; Mismatches 40; Indels 5; Gaps 1;
QY 1 MAKQYDVLFRLLIGDSGVGKTCILCRFTDNEPHSSHISTIGVDFKMKTIYVDGKIVRIQ 60
DB 1 MAKTYDYLFRLLIGDSGVGKTCILCRFTDNEPHSSHISTIGVDFKMKTIYVDGKIVRIQ 60
QY 61 IWDTAGOERYQTITTKOYIRRAOGIFLVYDISSERSYQHIMKWVSDYDEVAPEGVQKILIG 120
DB 61 IWDTAGOERFHTITTYIRGANGIMLVYDITNGKSFENIKWIRNIEENASADVERKMLLG 120
QY 121 NKADEQKROGVREGOQOLAKEYGMDFYETSACTNLNIKESFTRLTELVLQAHKLEGL 180
DB 121 NKCELTDKROVSKERGEQIAIEYGIKFMETSAKASINVEAF-----LTIASDIKARTEK 175
QY 181 RMRASN 186
DB 176 RMEANN 181
RESULT 14
ID AAG67154
XX AAG67154 standard; Protein; 207 AA.
XX AAG67154;
XX 13-NOV-2001 (first entry)
XX Amino acid sequence of human 27423 G-protein.
XX Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;

Search completed: November 17, 2002, 20:38:26
Job time : 68 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 18:05:49 ; Search time 92 Seconds

(without alignments)

13081.533 Million cell updates/sec

Title: US-09-817-198A-1

Perfect score: 3257

Sequence: 1 tgcgcgtgccgcgcgcag.....aaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 328367 seqs, 18475608 residues

Total number of hits satisfying chosen parameters: 656734

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	3257	100.0	3257	10	US-09-817-198A-1
2	2696.6	82.8	28770	10	US-09-817-198A-3
3	599	18.4	601	10	US-09-817-198A-28
4	599	18.4	601	10	US-09-817-198A-29
5	540	16.6	601	10	US-09-817-198A-30
6	516.6	15.9	601	10	US-09-817-198A-31
7	477.4	14.7	481	10	US-09-920-300A-303
8	477.4	14.7	481	12	US-10-033-528-303
9	438.8	13.5	463	10	US-09-964-824A-26
10	306.8	9.4	310	9	US-10-046-935-1930
11	217.8	6.7	624	10	US-09-794-257-9
12	217.8	6.7	1161	10	US-09-794-257-7
13	217.8	6.7	2497	10	US-09-834-975-879
14	217.8	6.7	2497	10	US-09-834-975-885
15	217.8	6.7	2497	10	US-09-834-975-894
16	217.8	6.7	2497	10	US-09-834-975-896
17	210	6.4	601	10	US-09-817-198A-32
18	198.4	6.1	1537	10	US-09-925-300-631
19	192	5.9	1274	10	US-09-925-302-91

20	173.4	5.3	651	9	US-09-938-842A-836	Sequence 836, App
21	156	4.8	925	10	US-09-967-736-4	Sequence 4, Appli
22	149.6	4.6	1673	9	US-09-954-531-1359	Sequence 1359, Ap
23	146	4.5	601	10	US-09-817-198A-33	Sequence 33, Appl
24	144.6	4.4	3936	10	US-09-919-172-49	Sequence 49, Appl
25	142.6	4.4	585	10	US-09-917-800A-1461	Sequence 1461, Ap
26	136.2	4.2	875	12	US-10-051-986-10	Sequence 10, Appl
27	136.2	4.2	1116	10	US-09-794-257-13	Sequence 13, Appl
28	136.2	4.2	1137	9	US-10-108-605-44	Sequence 44, Appl
29	136.2	4.2	2674	10	US-09-817-199A-1	Sequence 1, Appli
30	132.6	4.1	896	10	US-09-770-445-478	Sequence 478, App
31	131.8	4.0	576	10	US-09-794-257-15	Sequence 15, Appl
32	127.4	3.9	881	10	US-09-770-445-529	Sequence 529, App
33	124.2	3.8	4083	10	US-09-817-182-1	Sequence 1, Appli
34	123.6	3.8	1022	10	US-09-920-300A-1703	Sequence 1703, Ap
35	123.6	3.8	1022	12	US-10-033-528-1703	Sequence 1703, Ap
36	120.8	3.7	857	10	US-09-917-800A-1426	Sequence 1426, Ap
37	120.8	3.7	1129	10	US-09-925-301-235	Sequence 235, App
38	120.6	3.7	847	10	US-09-988-974-4	Sequence 4, Appli
39	120	3.7	654	9	US-09-938-842A-2113	Sequence 2113, Ap
40	119.4	3.7	609	9	US-09-938-842A-832	Sequence 832, App
41	115.2	3.5	676	10	US-09-770-149-374	Sequence 374, App
42	114.6	3.5	645	9	US-09-938-842A-774	Sequence 774, App
43	114.6	3.5	900	10	US-09-770-445-469	Sequence 469, App
44	113	3.5	639	10	US-09-350-874-66	Sequence 66, Appl
45	112	3.4	894	10	US-09-770-445-487	Sequence 487, App

ALIGNMENTS

RESULT 1

US-09-817-198A-1

; Sequence 1, Application US/09817198A

; Patent No. US20020146758A1

; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: CLO01188

; CURRENT APPLICATION NUMBER: US/09/817.198A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 3257

; TYPE: DNA

; ORGANISM: Human

US-09-817-198A-1

Query Match	100.0%	Score 3257;	DB 10;	Length 3257;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 3257;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Oy	1	TGCCCGCTGCCCGCCGAGTTC	CGCCCGCCGCTGGCCCGAGT	CGCCCGAGCAGTACG 60
Db	1	TGCCCGCTGCCCGCCGAGTTC	CGCCCGCCGCTGGCCCGAGC	CGCCCGAGCAGTACG 60
Oy	61	ATGTGCTGTTCGGGCTGTGCT	GGGAGTCTCGGGGTGGGCA	AGACCTGCCCTGCTGT 120
Db	61	ATGTGCTGTTCGGGCTGTGCT	GGGAGTCTCGGGGTGGGCA	AGACCTGCCCTGCTGT 120
Oy	121	GCCGCTTACCACCAACAGTTC	CACTCCTCGCATCTCCACC	ATCGGTGTGACTTTA 180
Db	121	GCCGCTTACCACCAACAGTTC	CACTCCTCGCATCTCCACC	ATCGGTGTGACTTTA 180
Oy	181	AGATGAAGACCATAGAGGTAG	ACGCATCAAAAGTCCGAT	ACAGATCTGGGACACTGCAG 240
Db	181	AGATGAAGACCATAGAGGTAG	ACGCATCAAAAGTCCGAT	ACAGATCTGGGACACTGCAG 240
Oy	241	GCCAGGAGATACACAGACCAT	CACAAAGCAGTACTATCG	CGCGGCCCGCCAGGGATATTTT 300

Db 241 GCGAGGAGATACAGACCATCACAAAGCAGTACTATCGCGGGGCCCCAGGGGATATTTT 300
QY 301 TGGTCTATGACATTAAGCAGGCGCTCTTACCAGCACATCATGAAGTGGGTGAGTGAG 360
Db 301 TGGTCTATGACATTAAGCAGGCGCTCTTACCAGCACATCATGAAGTGGGTGAGTGAG 360
QY 361 TGGATGAGTACGACAGAAAGGCTCCAGAAATCCTTATTTGGGAATAAGGCTGATGAGG 420
Db 361 TGGATGAGTACGACAGAAAGGCTCCAGAAATCCTTATTTGGGAATAAGGCTGATGAGG 420
QY 421 ACCAGAAACGGCAGGTGGGAAGAGCAGAGGCGCAGAGCTGGCGAAGTAGTATGCGATGG 480
Db 421 AGCAGAAACGGCAGGTGGGAAGAGCAGAGGCGCAGAGCTGGCGAAGTAGTATGCGATGG 480
QY 481 ACTTCTATGAAACAGTGCCTGCACCAACTCAACATTAAGAGATCATTCACGCGCTCA 540
Db 481 ACTTCTATGAAACAGTGCCTGCACCAACTCAACATTAAGAGATCATTCACGCGCTCA 540
QY 541 CAGAGCTGCTGTCAGGCCCCATAGGAAGAGCTGGAAGGCTCCGGATGCGTGCAGCA 600
Db 541 CAGAGCTGCTGTCAGGCCCCATAGGAAGAGCTGGAAGGCTCCGGATGCGTGCAGCA 600
QY 601 ATGAGTTGCACTGGCAGAGCTGGAGGAGGAGGAGGCAACCCAGAGGCCCCAGCGAACT 660
Db 601 ATGAGTTGCACTGGCAGAGCTGGAGGAGGAGGAGGCAACCCAGAGGCCCCAGCGAACT 660
QY 661 CTTTCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 CTTTCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 CTTCAAGAGGCGGTGGGAGACAGGGGAGCGGGGCTTTGCGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CTTCAAGAGGCGGTGGGAGACAGGGGAGCGGGGCTTTGCGCTGCTGCTGCTGCTGCTGCT 780
QY 781 GTGATCAGCCTATTTAGTATACATAGCCACTACTGCTCCCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 GTGATCAGCCTATTTAGTATACATAGCCACTACTGCTCCCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 CTGCTGCTATCTCAAGCAGCCCCGTGCCAGCCCCGTCCACCCCTGGAGTGGTCTTCTTCA 900
Db 841 CTGCTGCTATCTCAAGCAGCCCCGTGCCAGCCCCGTCCACCCCTGGAGTGGTCTTCTTCA 900
QY 901 GCCTGTTTCCAGCAGGCGCTGTAGACCCCCACAGATGTGCCGGAAGCACTGTCTC 960
Db 901 GCCTGTTTCCAGCAGGCGCTGTAGACCCCCACAGATGTGCCGGAAGCACTGTCTC 960
QY 961 ACCATCCGACCCACAGACAGCAGGCGTGAGTCCAGGCGCACCTTTACAGTGCCTC 1020
Db 961 ACCATCCGACCCACAGACAGCAGGCGTGAGTCCAGGCGCACCTTTACAGTGCCTC 1020
QY 1021 CTTTCTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 CTTTCTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 TGACCCCTCCCGTGGTGGTTCGTATCAAGCTCTCAACCCCGTCCCGCTGCTGCT 1140
Db 1081 TGACCCCTCCCGTGGTGGTTCGTATCAAGCTCTCAACCCCGTCCCGCTGCTGCTGCT 1140
QY 1141 CTTGCTGTGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 CTTGCTGTGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 GCTCGTGGGAGGTTCACCCCTTGGATCCAGGAAGAACCTCCACCCCTGCTCGTGGGTG 1260
Db 1201 GCTCGTGGGAGGTTCACCCCTTGGATCCAGGAAGAACCTCCACCCCTGCTCGTGGGTG 1260
QY 1261 GGCCAAAGGCTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db 1261 GGCCAAAGGCTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1321 TGGGCTGCTCCCTCAGTGCCTGCGAAAGTGGAGGATCGAGGTAGGAGGAAACAGCAA 1380
Db 1321 TGGGCTGCTCCCTCAGTGCCTGCGAAAGTGGAGGATCGAGGTAGGAGGAAACAGCAA 1380

QY 1381 CCGGGAGTCTCGAGCCTGGGGCTGCCCTACTCTACCCATTTCCCGACAGAGCTTTG 1440
Db 1381 CCGGGAGTCTCGAGCCTGGGGCTGCCCTACTCTACCCATTTCCCGACAGAGCTTTG 1440
QY 1441 CCTTTGCTTGGCTGCCCGCTGCTTTTGGGAACTGAGCTCAGAGGAGGTGCTTCAG 1500
Db 1441 CCTTTGCTTGGCTGCCCGCTGCTTTTGGGAACTGAGCTCAGAGGAGGTGCTTCAG 1500
QY 1501 AGAAGGAAACAAATGAGGGTGGAGGATAAAGTCAACCTCCATTTCTTACCTCCCA 1560
Db 1501 AGAAGGAAACAAATGAGGGTGGAGGATAAAGTCAACCTCCATTTCTTACCTCCCA 1560
QY 1561 TGCAGCATGAACACAAATTTCTCTCCACCTGGCTCCAAATTTAAAGATGTGAGCAAGGC 1620
Db 1561 TGCAGCATGAACACAAATTTCTCTCCACCTGGCTCCAAATTTAAAGATGTGAGCAAGGC 1620
QY 1621 CTGTGGGTACTCCAGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTGAGGCCAACCAT 1680
Db 1621 CTGTGGGTACTCCAGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTGAGGCCAACCAT 1680
QY 1681 GCACCTCACAAGGGGAGCATTTGGAATGAAGACTAGCTCCTATGTATCAGGTTAAGA 1740
Db 1681 GCACCTCACAAGGGGAGCATTTGGAATGAAGACTAGCTCCTATGTATCAGGTTAAGA 1740
QY 1741 GCAAGGGAGGCTGGCCAGGGACAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800
Db 1741 GCAAGGGAGGCTGGCCAGGGACAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800
QY 1801 GGGCTCTTAGGCCCTTAAAGCAAGCGCTCATCTTCTAGTAAGAGAGCATTTCTCAGACTCCCA 1860
Db 1801 GGGCTCTTAGGCCCTTAAAGCAAGCGCTCATCTTCTAGTAAGAGAGCATTTCTCAGACTCCCA 1860
QY 1861 GCGGAGGACTGAGCCTAGCCTTACCAACCAAGTTCTCCTGGGACCCAAAGTTTATGG 1920
Db 1861 GCGGAGGACTGAGCCTAGCCTTACCAACCAAGTTCTCCTGGGACCCAAAGTTTATGG 1920
QY 1921 GAGAAGGCAAGACATTCATGGGAAGAGAGAGGAGGCGCTGGGTAGAACCTTTGGTG 1980
Db 1921 GAGAAGGCAAGACATTCATGGGAAGAGAGAGGAGGCGCTGGGTAGAACCTTTGGTG 1980
QY 1981 CTGTTCTCTTGGCCTTAAAGCAAGCGCTCATCTTGGCCTTACTCTCTCTGATAGGCTT 2040
Db 1981 CTGTTCTCTTGGCCTTAAAGCAAGCGCTCATCTTGGCCTTACTCTCTCTGATAGGCTT 2040
QY 2041 GAGGTTTGCACACCACTGTGGCTACAGGTGAGGAGAGGAGTCTCTTCTCCAGAG 2100
Db 2041 GAGGTTTGCACACCACTGTGGCTACAGGTGAGGAGAGGAGTCTCTTCTCCAGAG 2100
QY 2101 TGTATGTTTCCAGGAAGTTTCTTAAACCCATATGGCCCAAGATAGCTCGTAGGAGCCC 2160
Db 2101 TGTATGTTTCCAGGAAGTTTCTTAAACCCATATGGCCCAAGATAGCTCGTAGGAGCCC 2160
QY 2161 TTTAAAGAGGAAACAGTAAATTTACCAGTTCTACTGGGTTCTGCCCACCGTCCCAAG 2220
Db 2161 TTTAAAGAGGAAACAGTAAATTTACCAGTTCTACTGGGTTCTGCCCACCGTCCCAAG 2220
QY 2221 TGGGCGAGGCTTAGGAAGGGTCAATTTAAAGCCACACATTAGCTGCAGTGGCTG 2280
Db 2221 TGGGCGAGGCTTAGGAAGGGTCAATTTAAAGCCACACATTAGCTGCAGTGGCTG 2280
QY 2281 CAGCCAAACAAAGACTGGGTGTAGTATTCATCACTAAGAACCAATCCAGGGCA 2340
Db 2281 CAGCCAAACAAAGACTGGGTGTAGTATTCATCACTAAGAACCAATCCAGGGCA 2340
QY 2341 CTATATGTCAAGGATAAGAACTCTTCTTACTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
Db 2341 CTATATGTCAAGGATAAGAACTCTTCTTACTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
QY 2401 ACCATCAAACTTCT 2460
Db 2401 ACCATCAAACTTCT 2460

Qy	2461	GGACTTAGGGACAGGGTATAGCTTATAGTGGTGGAAAGCAAAAGGAGAGACGAGGAAGTTGTAA	2520
Db	2461.	GGACTTTAGGGACAGGGTATAGCTTATAGTGGTGGAAAGCAAAAGGAGAGACGAGGAAGTTGTAA	2520
Qy	2521	ATCACATGGCTATATGAGAAAGCAGACACTAACCTCTAGAGTGAAGCTCTGCATAGGCTGG	2580
Db	2521.	ATCACATGGCTATATGAGAAAGCAGACACTAACCTCTAGAGTGAAGCTCTGCATAGGCTGG	2580
Qy	2581	AGTTGCTTCTTTGAAGATGGGACTCCTTTGGGTATCAAGACCTATGCCACATCACACTGGG	2640
Db	2581.	AGTTGCTTCTTTGAAGATGGGACTCCTTTGGGTATCAAGACCTATGCCACATCACACTGGG	2640
Qy	2641	GCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCACGACGGGACTTGAGAAAGTT	2700
Db	2641.	GCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCACGACGGGACTTGAGAAAGTT	2700
Qy	2701	ATATTGGGCAGTGGCTCCAAATCTGTGGACCACTATTTTCAGACTTTCCCTGAAGATCACGGCA	2760
Db	2701.	ATATTGGGCAGTGGCTCCAAATCTGTGGACCACTATTTTCAGACTTTCCCTGAAGATCACGGCA	2760
Qy	2761	GGGTGCCATTCAATGTCTTTCTCTCTAGCCCCCTCAGGAAAGAGCACTATATTTGTATC	2820
Db	2761.	GGGTGCCATTCAATGTCTTTCTCTCTAGCCCCCTCAGGAAAGAGCACTATATTTGTATC	2820
Qy	2821	TGTACCCCTAGGGGTTCTGGAAGGGAAAAACATGGAAATCAGGAATCTCTATAGACTGATAGGCC	2880
Db	2821.	TGTACCCCTAGGGGTTCTGGAAGGGAAAAACATGGAAATCAGGAATCTCTATAGACTGATAGGCC	2880
Qy	2881	CTATCCACAAGGGCCATGACTGGGAAAAAGGTATGGAGCACAAGAGAAATTTGGGATTTTA	2940
Db	2881.	CTATCCACAAGGGCCATGACTGGGAAAAAGGTATGGAGCACAAGAGAAATTTGGGATTTTA	2940
Qy	2941	GGGTGCAGCTACGCTCACCCCTAAACTTTTGGTGGCCCTGGGCGATGTCTTGAGGCCACAGAC	3000
Db	2941.	GGGTGCAGCTACGCTCACCCCTAAACTTTTGGTGGCCCTGGGCGATGTCTTGAGGCCACAGAC	3000
Qy	3001	TGTTTAAGCAGGCTCTCTGGCCGTGTTTACTCGTCACCACTCTGCACCTGCTGCTCTTGAG	3060
Db	3001.	TGTTTAAGCAGGCTCTCTGGCCGTGTTTACTCGTCACCACTCTGCACCTGCTGCTCTTGAG	3060
Qy	3061	ACTCCATCCAGCCCCCAGGCACGCCACCTGCTGCTCGAGCCCTCCACTATCTCCCTGTGACGG	3120
Db	3061.	ACTCCATCCAGCCCCCAGGCACGCCACCTGCTGCTCGAGCCCTCCACTATCTCCCTGTGACGG	3120
Qy	3121	GTGAACCTTGGTACTGTCTCTCGGGTCCATATATGAATTTGTGAGCAGGGTTCACTATT	3180
Db	3121.	GTGAACCTTGGTACTGTCTCTCGGGTCCATATATGAATTTGTGAGCAGGGTTCACTATT	3180
Qy	3181	TTAAACACAGATGTTTTACAAAAATAAGATTATTTTCAAACACCACCAAAAAA	3240
Db	3181.	TTAAACACAGATGTTTTACAAAAATAAGATTATTTTCAAACACCACCAAAAAA	3240
Qy	3241	AAAAAAAAAAAAAAAAA 3257	
Db	3241.	AAAAAAAAAAAAAAAAA 3257	

RESULT 2

```

RESOLUTION: 2.00
US-09-817-198A-3
; Sequence 3, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28770

```

Query Match	82.8%	Score 2696.6	DB 10	Length 28770
Best Local Similarity	98.8%	Pred. No. 0		
Matches 2717	Conservative 0	Mismatches 34	Indels 0	Gaps 0
QY 473	TGGGATGACATTCATGAAACAAGTGCCTGCACCAACCTCAACATTAAGAGTCATTCCAC	532		
DB 25562	TGCCAGGGAGAGGAGACACTGGACTACCTGTGCCCTTTGGTTCCAGTCATTCCAC	25621		
QY 533	CGCTGTACAGAGCTGTGCTGCAGGCCCATAGGAAGGAGCTGGAAGCCCTCCGGATGCG	592		
DB 25622	CGCTGTACAGAGCTGTGCTGCAGGCCCATAGGAAGGAGCTGGAAGCCCTCCGGATGCG	25681		
QY 593	TGCCAGCAATGAGTTGGCACTTGGCAGACTGGAGGAGGAGGAGGCAACCCGAGGGCCC	652		
DB 25682	TGCCAGCAATGAGTTGGCACTTGGCAGACTGGAGGAGGAGGAGGCAACCCGAGGGCCC	25741		
QY 653	AGCGAATCTTTCGAAACCTCTGCTGTGCTGAGTCTCTGTGTGGGCACCCACACGACACC	712		
DB 25742	AGCGAATCTTTCGAAACCTCTGCTGTGCTGAGTCTCTGTGTGGGCACCCACACGACACC	25801		
QY 713	CCTCTTCCTCAGGAGGCCCTGGGCGACAGGAGGAGCGGGGGCTTTGCCCTGTGCTGT	772		
DB 25802	CCTCTTCCTCAGGAGGCCCTGGGCGACAGGAGGAGCGGGGGCTTTGCCCTGTGCTGT	25861		
QY 773	CCTCTGCTGTGATGACCCATTTGAGTATTCAGTAGCCACTACTCCCCCTGCCTGGCCCTGA	832		
DB 25862	CCTCTGCTGTGATGACCCATTTGAGTATTCAGTAGCCACTACTCCCCCTGCCTGGCCCTGA	25921		
QY 833	GAGGGCTCTGCTGTTCATCTCAAGCAGCCCTGTCTCCAGCCCGCTGCCACCTGGAGTGGT	892		
DB 25922	GAGGGCTCTGCTGTTCATCTCAAGCAGCCCTGTCTCCAGCCCGCTGCCACCTGGAGTGGT	25981		
QY 893	CTTCTTCAGCCTGTTTCCCGACCCACAGGCGCTGTACGACCCCGACAGATGTCGCGAAGC	952		
DB 25982	CTTCTTCAGCCTGTTTCCCGACCCACAGGCGCTGTACGACCCCGACAGATGTCGCGAAGC	26041		
QY 953	ACTGTCTACCATATCCCGCACCCACAGACACAGCCAGGGCTGGAGTCCAGGCCACTTTC	1012		
DB 26042	ACTGTCTACCATATCCCGCACCCACAGACACAGCCAGGGCTGGAGTCCAGGCCACTTTC	26101		
QY 1013	AGCTGCTCCTTTCTCCGTCATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1072		
DB 26102	AGCTGCTCCTTTCTCCGTCATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	26161		
QY 1073	TCTTTCTCTGACCCCTCCCTCCGCTCGGTTTCGTATCAAAAGCTCTCAAAACCCGTCCTC	1132		
DB 26162	TCTTTCTCTGACCCCTCCCTCCGCTCGGTTTCGTATCAAAAGCTCTCAAAACCCGTCCTC	26221		
QY 1133	CCGTGTCTCTGCTGTGTGCACTGCTCTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCT	1192		
DB 26222	CCGTGTCTCTGCTGTGTGCACTGCTCTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCT	26281		
QY 1193	GGACCCAGGCTCTGGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCCACCCCTGCCT	1252		
DB 26282	GGACCCAGGCTCTGGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCCACCCCTGCCT	26341		
QY 1253	CGTGGGTGGGCCAAGGCTACAGGGTGTCTTCTCTCTTCCCGCCACCCCTACTGTCCCTC	1312		
DB 26342	CGTGGGTGGGCCAAGGCTACAGGGTGTCTTCTCTCTTCCCGCCACCCCTACTGTCCCTC	26401		
QY 1313	ATGTGCCAATGGGCTGTCTCCCGAGTACCTGTGGAAAGTGAGCATGAGGTAGAGGGA	1372		
DB 26402	ATGTGCCAATGGGCTGTCTCCCGAGTACCTGTGGAAAGTGAGCATGAGGTAGAGGGA	26461		
QY 1373	AACAGCAACGGGAGTCTCTGAGCCCTGGGCTGCCCTACCTCTACCTTACCTTCCCGACCA	1432		
DB 26462	AACAGCAACGGGAGTCTCTGAGCCCTGGGCTGCCCTACCTCTACCTTACCTTCCCGACCA	26521		
QY 1433	GAGCTTTGCCCTTGTCTTGGCTGCCCGCTGCTCTTTTGGGGAACCTGAGCTCAGAGGCGAGG	1492		

Db 26522 GAGCTTGGCCCTTGCCTGGCTGCCGCCCTGCCCTTTGGGGAACCTGAGCTCAGAGCAGG 26581
QY 1493 TGCCTCAGAGGAAGAAACAAATAGGGGTGGCAGGGATATAAAGTACACCTCCATTCTCT 1552
Db 26582 TGCCTCAGAGGAAGAAACAAATAGGGGTGGCAGGGATATAAAGTACACCTCCATTCTCT 26641
QY 1553 ACCTCCCATGCAGCATGAACACAATTTCTCTCCACCTGGCTCCCAATTTAAAGATGTGG 1612
Db 26642 ACCTCCCATGCAGCATGAACACAATTTCTCTCCACCTGGCTCCCAATTTAAAGATGTGG 26701
QY 1613 ACCAAGCCCTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGTCAGTGACACTGTCAGG 1672
Db 26702 ACCAAGCCCTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGTCAGTGACACTGTCAGG 26761
QY 1673 CCAACATGCACATCCCAAGGGGAGCATTTGGAATTAAGAGTACTAGCTCCTATGTATCA 1732
Db 26762 CCAACATGCACATCCCAAGGGGAGCATTTGGAATTAAGAGTACTAGCTCCTATGTATCA 26821
QY 1733 GGTAAAGCAAGGAGAGCTGGCCAGGAGACAGCTTTGCACAGCAGAGGGGAATGTAG 1792
Db 26822 GGTAAAGCAAGGAGAGCTGGCCAGGAGACAGCTTTGCACAGCAGAGGGGAATGTAG 26881
QY 1793 CAACAGAGGCCCTCCTAGGCCCATCTTCCATTTCTTAGTAAGAGACATTTCTCTCA 1852
Db 26882 CAACAGAGGCCCTCCTAGGCCCATCTTCCATTTCTTAGTAAGAGACATTTCTCTCA 26941
QY 1853 GACTCCAGCGGAGGACTCAGGCTAGCCTTACGACCAACCAAGTTCCTCCGGGACCCAAA 1912
Db 26942 GACTCCAGCGGAGGACTCAGGCTAGCCTTACGACCAACCAAGTTCCTCCGGGACCCAAA 27001
QY 1913 GTTTATGGGAGAGGGCAAGACTTCATGGGAAGAGAGAGGAGGAGGAGGAGGAGGAGG 1972
Db 27002 GTTTATGGGAGAGGGCAAGACTTCATGGGAAGAGAGAGGAGGAGGAGGAGGAGGAGG 27061
QY 1973 GCTTGGTGTCTTCTTGGCTTTAAGACAAGCGCTCATCTGGCCCTTACCTCCTG 2032
Db 27062 GCTTGGTGTCTTCTTGGCTTTAAGACAAGCGCTCATCTGGCCCTTACCTCCTG 27121
QY 2033 ATAGGCTTGAAGGTTTCCCAACCACTGTGGCTACAGGTGGAGGGAAGAGGACTCCCTTC 2092
Db 27122 ATAGGCTTGAAGGTTTCCCAACCACTGTGGCTACAGGTGGAGGGAAGAGGACTCCCTTC 27181
QY 2093 CTCAGAGTCTATGTCAGAGAGTTCTTTAACCCCATATGGCCCAAGAGTACTCGTA 2152
Db 27182 CTCAGAGTCTATGTCAGAGAGTTCTTTAACCCCATATGGCCCAAGAGTACTCGTA 27241
QY 2153 GGAGGCCCTTTAAGAGCGGAACAGTAATTTACAGTTCTACTGGGTTCTCTGCCACCG 2212
Db 27242 GGAGGCCCTTTAAGAGCGGAACAGTAATTTACAGTTCTACTGGGTTCTCTGCCACCG 27301
QY 2213 TCCCAAGGTGGCGAGGCTTAGGAAGAGGCTCATTTTAAGCCACACATTTAGCTGCACTG 2272
Db 27302 TCCCAAGGTGGCGAGGCTTAGGAAGAGGCTCATTTTAAGCCACACATTTAGCTGCACTG 27361
QY 2273 CGTGGCTGAGCCAAAACAAGAACTGGGTGTGAGTATTCATCACTAAGAACCAAAAT 2332
Db 27362 CGTGGCTGAGCCAAAACAAGAACTGGGTGTGAGTATTCATCACTAAGAACCAAAAT 27421
QY 2333 CCAGGGCAGTCATATGTAAGGATAAGCACTTCTACTTCTCTCTCCAAAAGAGTG 2392
Db 27422 CCAGGGCAGTCATATGTAAGGATAAGCACTTCTACTTCTCTCTCCAAAAGAGTG 27481
QY 2393 GGGAAAGAACCATCAAACTTCTCTCTGACTTACCAAAACAGGAAACAGCAGGAGG 2452
Db 27482 GGGAAAGAACCATCAAACTTCTCTCTGACTTACCAAAACAGGAAACAGCAGGAGG 27541
QY 2453 GTGGCTCAGGACTTAGGAGCAGGTATAGCTTAGATGGTGAAGCAAGAGCAGGAGGA 2512
Db 27542 GTGGCTCAGGACTTAGGAGCAGGTATAGCTTAGATGGTGAAGCAAGAGCAGGAGGA 27601
QY 2513 AGTTGTAATACACTGGCTTAATGAGAAAGGAGACAGCTAACTCTAGGATGAAGCTGTGAC 2572

Db 27602 AGTTGTAATACACTGGCTAATGAGAAAGAGAGACGCTAACTCTAGATGAAGCTGTGAC 27661
QY 2573 TAGGCTGGAGTCTCTTCCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 2632
Db 27662 TAGGCTGGAGTCTCTTCCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 27721
QY 2633 AACTGGGCTTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTCTTTCACCCAGGAGCTT 2692
Db 27722 AACTGGGCTTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTCTTTCACCCAGGAGCTT 27781
QY 2693 GAGAAGTTATATTGGGAGTGGCTCCAATCTGTGGACAGTATTTCAAGCTTTCCCTGAAG 2752
Db 27782 GAGAAGTTATATTGGGAGTGGCTCCAATCTGTGGACAGTATTTCAAGCTTTCCCTGAAG 27841
QY 2753 ATCAGCAGGCTGCCATTCATTCTTCTCTAGCCCCCTCAGAAAGAGGACTAT 2812
Db 27842 ATCAGCAGGCTGCCATTCATTCTTCTCTAGCCCCCTCAGAAAGAGGACTAT 27901
QY 2813 ATTTGTACTGTACCTAGGGTTCTGGAAGGGAAACATGGAATCAGGATCTATAGACT 2872
Db 27902 ATTTGTACTGTACCTAGGGTTCTGGAAGGGAAACATGGAATCAGGATCTATAGACT 27961
QY 2873 GATAGGCCCTTATCCACAAGGGCCATGACTGGGAAAGGTATGGAGCAGAGGAGAAATG 2932
Db 27962 GATAGGCCCTTATCCACAAGGGCCATGACTGGGAAAGGTATGGAGCAGAGGAGAAATG 28021
QY 2933 GGATTTAGGCTGCAGCTACCTAACTTTTGGTGGCTGGGGCATGTCTTTGAG 2992
Db 28022 GGATTTAGGCTGCAGCTACCTAACTTTTGGTGGCTGGGGCATGTCTTTGAG 28081
QY 2993 GCCCAGACTTTAAGCAGGCTCTGTGGCTGTGTCTCTCCTCACCACCTCTGCACCTGCT 3052
Db 28082 GCCCAGACTTTAAGCAGGCTCTGTGGCTGTGTCTCTCCTCACCACCTCTGCACCTGCT 28141
QY 3053 GTCTTGAGACTCCATCCAGCCCGCAGGCACGCTCTCTGAGCCTCCACCTATCTCCC 3112
Db 28142 GTCTTGAGACTCCATCCAGCCCGCAGGCACGCTCTCTGAGCCTCCACCTATCTCCC 28201
QY 3113 TGTGAGGGTGAATCTGCTACTGTCTCGGGTCCCATATATGAATGTGAGCAGGCTT 3172
Db 28202 TGTGAGGGTGAATCTGCTACTGTCTCGGGTCCCATATATGAATGTGAGCAGGCTT 28261
QY 3173 CATCTATTTAAACACAGATGTTTACAAAATAAGATTTTCAACACCC 3223
Db 28262 CATCTATTTAAACACAGATGTTTACAAAATAAGATTTTCAACACCC 28312

RESULT 3

US-09-817-198A-28
; Sequence 28, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; PROTEINS, AND USES THEREOF
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-817-198A-28

Query Match 18.4%; Score 599; DB 10; Length 601;
Best Local Similarity 99.7%; Pred. No. 1.le-116;
Matches 599; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1076 TTCTCTGACCCCTCCGCTCGGTCGTTTCGTATCAAGCTCTCAAAACCCGCTCCCGG 1135

Db 121 AGAAGGACTATATTTGTACTGTACCTAGGGTTCTGGAAGGAAAAACATGAATCAGGA 180
QY 2862 TTCCTATAGACTGATAGGCCCTATCCACAGGCCCATGACTGGAAAAAGGTATGGAGCAG 2921
Db 181 TTCCTATAGACTGATAGGCCCTATCCACAGAGGCCCATGACTGGAAAAAGGTATGGAGCAG 240
QY 2922 AAGGAGAATTGGGATTTTAGGTGTCAGCTACGCTCACCTTAACCTTTTGGTGGCCCTGGGG 2981
Db 241 AAGGAGAATTGGGATTTTAGGTGTCAGCTACGCTCACCTTAACCTTTTGGTGGCCCTGGGG 300
QY 2982 CATGCTCTTGAGGCCAGACTGTTAAGCAGGCTCTGCTGGCCCTGTTTACTGTCACACCT 3041
Db 301 YATGCTCTTGAGGCCAGACTGTTAACCAGGCTCTGCTGGCCCTGTTTACTGTCACACCT 360
QY 3042 CTGACCTGCTGTCTTGAGACTCCATCCAGCCCCAGGCCACCTGCTCTGTAGCCCTC 3101
Db 361 CTGACCTGCTGTCTTGAGACTCCATCCAGCCCCAGGCCACCTGCTCTGTAGCCCTC 420
QY 3102 CACTATCTCCCTGTGACGGGTGAACCTTCTGTACTGTCTCGGTCCATATATGAATG 3161
Db 421 CACTATCTCCCTGTGACGGGTGAACCTTCTGTACTGTCTCGGTCCATATATGAATG 480
QY 3162 TGACGAGGTTTCATCTATTTTAAACACAGATGTTTACAAAAATAAGATTATTTCAACCA 3221
Db 481 TGACGAGGTTTCATCTATTTTAAACACAGATGTTTACAAAAATAAGATTATTTCAACCA 540
QY 3222 CC 3223
Db 541 CC 542

RESULT 6
US-09-817-198A-31
; Sequence 31, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-817-198A-31

Query Match 15.9%; Score 516.6; DB 10; Length 601;
Best Local Similarity 99.8%; Pred. No. 2.1e-99;
Matches 516; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2707 GGCAGTGGCTCCAACTGTGTGACACAGTATTTACAGCTTTCCCTGAAGATCAGGCGGTGC 2766
Db 1 GGCAGTGGCTCCAACTGTGTGACACAGTATTTACAGCTTTCCCTGAAGATCAGGCGGTGC 60
QY 2767 CATTCATTGTCTTCTCTCTAGCCCCCTCAGGAAGAAGGACTATATTTGTACTGTACC 2826
Db 61 CATTCATTGTCTTCTCTCTAGCCCCCTCAGGAAGAAGGACTATATTTGTACTGTACC 120
QY 2827 CTAGGGGTTCTGGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGGCCCTATCC 2886
Db 121 CTAGGGGTTCTGGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGGCCCTATCC 180
QY 2887 ACAAGGCCATGACTGGGAAAAGTATGCGACAGAGAGAAATTTAGGGTGC 2946
Db 181 ACAAGGCCATGACTGGGAAAAGTATGCGACAGAGAGAAATTTAGGGTGC 240
QY 2947 AGCTACGCTACCTAACTTTTGGTGGCCCTGGGCAATGCTTCTGAGGCCACAGACTGTAA 3006

Db 241 AGCTACGCTACCCCTAAACATTTTGGTGGCCCTGGGCGCATGCTTTGAGGCCACAGACTGTAA 300
QY 3007 GCAGGCTCTGCTGGCCTGTTTACTCGTCACACACCTCTGCACCTCTGCTCTTTGAGACTCCA 3066
Db 301 SCAGGCTCTGCTGGCCTGTTTACTCGTCACACACCTCTGCACCTCTGCTCTTTGAGACTCCA 360
QY 3067 TCCAGCCCCAGGACGCGCACCTGCTCTCGTGGCCCTCCACTATCTCCCTGTGAGGGGTGAAC 3126
Db 361 TCCAGCCCCAGGACGCGCACCTGCTCTCGTGGCCCTCCACTATCTCCCTGTGAGGGGTGAAC 420
QY 3127 TTCGTGCTACTGCTCGGTCCATATATGAATTTGTGAGCAGGTTTCATCTATTTTAAAC 3186
Db 421 TTCGTGCTACTGCTCGGTCCATATATGAATTTGTGAGCAGGTTTCATCTATTTTAAAC 480
QY 3187 ACAGATGTTTACAAATAAAGATTATTTCAAAACCACC 3223
Db 481 ACAGATGTTTACAAATAAAGATTATTTCAAAACCACC 517

RESULT 7
US-09-920-300A-303
; Sequence 303, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 417..461
; OTHER INFORMATION: n = A,T,C or G
US-09-920-300A-303

Query Match 14.7%; Score 477.4; DB 10; Length 481;
Best Local Similarity 99.4%; Pred. No. 3.2e-91;
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2582 GTTGCTTCTCTTGAAGATGGGACTCCCTTGGGTATCAAGACCTATGCCACATCAGCTGGG 2641
Db 1 GTTGCTTCTCTTGAAGATGGGACTCCCTTGGGTATCAAGACCTATGCCACATCAGCTGGG 60
QY 2642 CTAGGGAAGTAGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGAGCTTGAGAAGTTA 2701
Db 61 CTAGGGAAGTAGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGAGCTTGAGAAGTTA 120
QY 2702 TATGGGCACTGGCTCCATCTGTGGACCAAGTATTCAGCTTCCCTGAAGTACGCGAG 2761
Db 121 TATGGGCACTGGCTCCATCTGTGGACCAAGTATTCAGCTTCCCTGAAGTACGCGAG 180
QY 2762 GGTGCCATTATGCTCTTCTCTAGCCCCCTCAGGAAGAAGGACTATATTTTACT 2821
Db 181 GGTGCCATTATGCTCTTCTCTAGCCCCCTCAGGAAGAAGGACTATATTTTACT 240
QY 2822 GTACCTTAGGGTTCTGGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGCCC 2881
Db 241 GTACCTTAGGGTTCTGGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGCCC 300
QY 2882 TATCCACAGGGCCATGACTGGGAAAAGTATGCGACAGAGAGAAATTTGGGATTTAG 2941
Db 301 TATCCACAGGGCCATGACTGGGAAAAGTATGCGACAGAGAGAAATTTGGGATTTAG 360

QY 2942 GGTGAGCTAGCTCACCTAAACTTTTGGTGGCTGGGGCATGCTTTGAGGCCAGACT 3001
Db |||||||
Db 361 GGTGAGCTAGCTCACCTAAACTTTTGGTGGCTGGGGCATGCTTTGAGGCCANACT 420
QY 3002 GTTAAGCAGGCTGCTGGCTGTTTACTCGTCACCACTCTGCACCTGCTGCTTTGAGA 3061
Db |||||||
Db 421 GTTAACCAGGCTGCTGGCTGTTTACTCGTCACCACTCTGCACCTGCTGCTTTGAGA 480
QY 3062 C 3062
Db |
Db 481 C 481
RESULT 8
US-10-033-528-303
; Sequence 303, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 417, 461
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-303

Query Match 14.7%; Score 477.4; DB 12; Length 481;
Best Local Similarity 99.4%; Pred. No. 3.2e-91;
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2582 GTTGCTTCCTTGAAGATGGGACTCCTGGGTATCAAGACCTATGCCATCACATCACACTGGGG 2641
Db |||||||
Db 1 GTTGCTTCCTTGAAGATGGGACTCCTGGGTATCAAGACCTATGCCATCACACTGGGG 60
QY 2642 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGCTTTCAGCCAGGAGCTTGAGAAGTTA 2701
Db |||||||
Db 61 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGCTTTCAGCCAGGAGCTTGAGAAGTTA 120
QY 2702 TATTGGGCAGTGGCTCCAACTCTGGACCACTATTTAGCTTTCCCTGAAGATCAGGCAG 2761
Db |||||||
Db 121 TATTGGGCAGTGGCTCCAACTCTGGACCACTATTTAGCTTTCCCTGAAGATCAGGCAG 180
QY 2762 GTGTCATTATGTTCTTCTCTCTAGCCCTCAGGAAGAAGSACTATATTTGTACT 2821
Db |||||||
Db 181 GTGTCATTATGTTCTTCTCTCTAGCCCTCAGGAAGAAGSACTATATTTGTACT 240
QY 2822 GTACCTTAGGGGTTCTGGAAGGAAACATGGAATCAGGATTCATAGACTGATAGGCC 2881
Db |||||||
Db 241 GTACCTTAGGGGTTCTGGAAGGAAACATGGAATCAGGATTCATAGACTGATAGGCC 300
QY 2882 TATCCACAGGGCCATGACTGGGAAAGGATGAGGAGCAGAGGAATGCGGATTTTAG 2941
Db |||||||
Db 301 TATCCACAGGGCCATGACTGGGAAAGGATGAGGAGCAGAGGAATGCGGATTTTAG 360
QY 2942 GTGTCAGCTACGCTCACCTAACTTTTGGTGGCTGGGGCATGCTTTGAGGCCAGACT 3001
Db |||||||
Db 361 GTGTCAGCTACGCTCACCTAACTTTTGGTGGCTGGGGCATGCTTTGAGGCCANACT 420
QY 3002 GTTAAGCAGGCTGCTGGCTGTTTACTCGTCACCACTCTGCACCTGCTGCTTTGAGA 3061
Db |||||||

Db 421 GTTAACCAGGCTGCTGGCTGTTTACTCGTCACCACTCTGCACCTGCTGCTTTGAGA 480
QY 3062 C 3062
Db |
Db 481 C 481
RESULT 9
US-09-964-824A-26/c
; Sequence 26, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-26

Query Match 13.5%; Score 438.8; DB 10; Length 463;
Best Local Similarity 99.3%; Pred. No. 3.9e-83;
Matches 451; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 2790 CCCCTCAGGAAAGAGACTATATTTGACTGTACCTTAGCGGTTCTGGAAGGAAAAAC 2849
Db |||||||
Db 453 CCCCTCAGGAAAGAGACTATATTTGACTGTACCTTAGCGGTTCTGGAAGGAAAAAC 394
QY 2850 ATGGAATCAGGATTCATAGACTGATAGCCCTATCCACAAGGGCATGACTGGGAAAG 2909
Db |||||||
Db 393 ATGGAATCAGGATTCATAGACTGATAGCCCTATCCACAAGGGCATGACTGGGAAAG 334
QY 2910 GTATGGAGCAGAGGAATTTGGATTTTAGGTGCGAGCTACGCTACCCCTAAACTTTT 2969
Db |||||||
Db 333 GTATGGAGCAGAGGAATTTGGATTTTAGGGTGCAG-TACGCTCACCCCTAAACTTTT 275
QY 2970 GGTGCTCTGGGTCATGCTTTGAGGCCAGACTGTTTAAAGCAGGCTCTGCTGGCTGTTTAC 3029
Db |||||||
Db 274 GGTGCTCTGGGTCATGCTTTGAGGCCAGACTGTTTAAAGCAGGCTCTGCTGGCTGTTTAC 215
QY 3030 TGTCCACCACTCTGCACCTGCTGTTTGAGACTCCATCCAGCCCCCAGCAGCCACCTG 3089
Db |||||||
Db 214 TGTCCACCACTCTGCACCTGCTGTTTGAGACTCCATCCAGCCCCCAGCAGCCACCTG 155
QY 3090 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAATTCGTGTACTGTCTCGGGTCC 3149
Db |||||||
Db 154 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAATTCGTGTACTGTCTCGGGTCC 95
QY 3150 ATATATGAATTCGTGAGCAGGTTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGAT 3209
Db |||||||
Db 94 ATATATGAATTCGTGAGCAGGTTTCATCTATTTTAAACACAGATGTTTACAAAATAAAGAT 35
QY 3210 TATTTTCAACCAACCAAAAAA 3243
Db |||||||
Db 34 TATTTTCAACCAACCAAAAAA 1

RESULT 10
US-10-046-935-1930
; Sequence 1930, Application US/10046935
; Patent No. US20020156011A1

GENERAL INFORMATION:
APPLICANT: Jiang, Yuqiu
APPLICANT: Harlocker, Susan L.
APPLICANT: Secrist, Heather
APPLICANT: Wang, Aijun
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.527C1
CURRENT APPLICATION NUMBER: US/10/046,935
CURRENT FILING DATE: 2002-01-15
NUMBER OF SEQ ID NOS: 2239
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1930
LENGTH: 310
TYPE: DNA
ORGANISM: Homo sapiens
US-10-046-935-1930

Query Match 9.4%; Score 306.8; DB 9; Length 310;
Best Local Similarity 99.4%; Pred. No. 1.6e-55;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2824 ACCTAGGGTCTTGAAGGAAACATGAATCAGATTCTATAGACTGATAGGCCCTA 2883
Db 1 ACCTAGGGTCTTGAAGGAAACATGAATCAGATTCTATAGACTGATAGGCCCTA 60
QY 2884 TCCACAGGGCCATGACTGGGAAAGTATGGGAGAGGAGAAATGGGATTTAGGG 2943
Db 61 TCCACAGGGCCATGACTGGGAAAGTATGGGAGAGGAGAAATGGGATTTAGGG 120
QY 2944 TGAGCTACCTCACCTAACTTTTGGTGGCTGGGCGATGCTTGGAGCCCGAGACTGT 3003
Db 121 TGAGCTACCTCACCTAACTTTTGGTGGCTGGGCGATGCTTGGAGCCCGAGACTGT 180
QY 3004 TAAGCAGGCTCTGTGGCCCTTTACTGCTGCACACCTCTGCACCTGCTCTTGAGACT 3063
Db 181 TAAGCAGGCTCTGTGGCCCTTTACTGCTGCACACCTCTGCACCTGCTCTTGAGACT 240
QY 3064 CCATCCAGCCCCAGGCGACCTGCTGCTGAGCCCTCCACTATCTCCCTGTGACGGGTG 3123
Db 241 CCATCCAGCCCCAGGCGACCTGCTGCTGAGCCCTCCACTATCTCCCTGTGACGGGTG 300

RESULT 11
US-09-794-257-9
; Sequence 9, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794, 257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 624
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-794-257-9

Query Match 6.7%; Score 217.8; DB 10; Length 624;
Best Local Similarity 65.6%; Pred. No. 9e-37;
Matches 318; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 45 ATGCGAAGCAGTACGATGTGCTTCCGGCTGTGCTGATCGGGGACTCCGGGGTGGGC 104
Db 1 ATGCGAAGCAGTACGATGTGCTTCAAGCTCCTGCTGATCGGGGACTCCGGGGTGGGC 60
QY 105 AAGACCTGCCTGCTGCGGCTTACCGACACGAGTTCCACTCCTCGCATCTCCACC 164
Db 61 AAGACCTGCCTGCTGCGGCTTCTCAGAGGACGCTTCAACACCACCTTCACTCTCCACC 120
QY 165 ATCGGTGTGACTTTAAGATGAAGACCATATAGAGTAGGTCAGGCGATCAAAAGTCCGGATACAG 224
Db 121 ATCGGAATGATTTTAAATTTAGACGATAGAACTAGATGGAAGAAATTAAGCTTCAAG 180
QY 225 ATCTGGACACTGCGAGGCGAGAGATACAGACCATCAACAAAGCAGTACTATCGGGCG 284
Db 181 ATATGGACACAGCGGGTTCAGGAAAGATTCGGAACAAATCACGACGCTACTTACAGAGA 240
QY 285 GCCAGGGGATATTTTGGTCTATGACATTTAGCAGGAGCGCTCTTACCAGCACATCATG 344
Db 241 GCATGGGCATTTATGCTGCTATGACATCACAATGAAATCTTTGCACATATTTAA 300
QY 345 AAGTGGTCTGAGTACGTCAGTACGACCCAGACAGGCGTCCAGAAATCTTATTTGGG 404
Db 301 AATGGATCAGAAACATTTGAAGAGCATGCTCTTCCGATGTCGAAAGATGATCTGGGT 360
QY 405 AATAAGCTGATGAGGAGCAGAAACGGCAGGTGGGAAGAGCAAGGCGAGCGCTGGCG 464
Db 361 AACAATGTGATGATGAATGACAAAGACAAAGTGTCAAAAGAAAGAGGGGAGAGCTAGCA 420
QY 465 AAGGAGTATGGATGACTTCTATGAACAAGTGTCTGCACCAACCTCAACATTAAGAG 524
Db 421 ATTGACTATGGATTAATTTTGGAGACAAGCGCAAAATCCAGTCAATATAGAGAG 480
QY 525 TCATT 529
Db 481 GCATT 485

RESULT 12
US-09-794-257-7
; Sequence 7, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1el
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794, 257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)...(641)
US-09-794-257-7

Query Match 6.7%; Score 217.8; DB 10; Length 1161;
Best Local Similarity 65.6%; Pred. No. 1.1e-36;
Matches 318; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 45 ATGCGAAGCAGTACGATGTGCTTCCGGCTGTGCTGATCGGGGACTCCGGGGTGGGC 104
Db 18 ATGCGAAGCAGTACGATGTGCTTCCAGCTCTGCTGATCGGGGACTCCGGGGTGGGC 77
QY 105 AAGACCTGCCTGCTGCGGCTTACCGACCAAGAGTTCCACTCCTCGCACATCTCCACC 164
Db 78 AAGACCTGCCTGCTGCGGCTTCTCAGAGGAGCGCTTCAACACCACCTTCACTCTCCACC 137

QY 165 ATCGGTGTTGACTTTAAGATGAAGACCATAGAGGTAGCGCATCAAAAGTCGGGATACAG 224
Db 138 ATCGGAATTGATTTTAAATTTAGAACGATAGAACTAGATGGAAGAAAATTAAGCTTCAG 197
QY 225 ATCTGGGACATGCGAGGACGAGAGATACCAGACCATCAAAAGCACTACTATCGGGCGG 284
Db 198 ATATGGGACACAGCGGGTCAGGAAGATTCCGAACAATCAGCAGCGCTACTACAGAGGA 257
QY 285 GCCAGGGGATATTTTGGTCTATGACATTTAGCAGCGAGCGCTCTTACCAGCACATCATG 344
Db 258 GCCATGGGCATTTATGCTGTCTATGACATCACAATGAAATTCCTTTGACAAATTTAAA 317
QY 345 AAGTGGGTGAGTGCAGTGCATGAGTACGACCAAGCGCTCCAGAAAGATCCTTATTGGG 404
Db 318 AATTGGATCAGAAACATTTGAAGAGCATGCTCTTCCGATGTCGAAAGAATGATCCTGGGT 377
QY 405 AATAGGCTGATGAGGACAGAAAGCGGAGGTGGGAAAGAGAGGAGCGGACGAGTGGCG 464
Db 378 AACAATGTGATATGAATGACAAAAGCAAGTGTCAAAAGAAAGAGGAGGAGAACTAGCA 437
QY 465 AAGGATATGGCATGGACTTCTATGAAACAAGTGCCTGCACCAACCTCAACATTTAAAGAG 524
Db 438 ATGACTATGGGATTAATTTCTTGGAGACAAGCGCAAAATCCAGTGCATGTAAGAGAG 497
QY 525 TCATT 529
Db 498 GCATT 502

RESULT 13
US-09-834-975-879
; Sequence 879, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 879
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-879

Query Match 6.7%; Score 217.8; DB 10; Length 2497;
Best Local Similarity 65.6%; Pred. No. 1.5e-36;
Matches 318; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 45 ATGGCGAAGCAGTACGATGTGCTTCCGGCTGCTGCTGATCGGGGACTCCGGGTGGCG 104
Db 92 ATGGCGAAGCAGTACGATTTATCTTCAAGCTCTCTGATCGGGGACTCGGGGTAGGC 151
QY 105 AAGACTGCCTGCTGCGGCTTCAACGACACAGAGTTCCTCTCCGACATCTCCACC 164
Db 152 AAGACTGCCCTGCTTCCGCTTCTCAGAGGCGCTTCAACACCACTTCATCTCCACC 211
QY 165 ATCGGTGTTGACTTTAAGATGAAGACCATAGAGGTAGCGCATCAAAAGTCGGGATACAG 224
Db 212 ATCGGAATTGATTTTAAATTTAGAACGATAGAACTAGATGGAAGAAAATTAAGCTTCAG 271

QY 225 ATCTGGGACACTGAGGCGAGGAGATACACAGACCATCAAAAGCAGTACTACTATCGGGCGG 284
Db 272 ATATGGGACACAGCGGGTCAGGAAGATTCCGAAACAATCACGACGCGTACTACAGAGGA 331
QY 285 GCCAGGGGATATTTTGGTCTATGACATTTAGCAGCGAGCGCTCTTACCAGCACATCATG 344
Db 332 GCCATGGGCATTTATGCTGTCTATGACATTCACAAATGAAAAATCCTTTGACAATATAAA 391
QY 345 AAGTGGGTGAGTGCAGTGCATGAGTACGACCAAGCGCTCCAGAAAGATCCTTATTGGG 404
Db 392 AATTGGATCAGAAACATTTGAAGAGCATGCTCTTCCGATGTCGAAAGAATGATCTGGGT 451
QY 405 AATAAGGCTGATGAGGACAGAAAGCGGAGGTGGGAAAGAGACAAGGCGAGCGAGTGGCG 464
Db 452 RACAATGTGATATGAATGACAAAAGACAAGTGTCAAAAGAAAGAGGAGAGCTAGCA 511
QY 465 AAGGATATGGCATGGACTTCTATGAAACAAGTGCCTGCACCACTCAACATTTAAAGAG 524
Db 512 ATTGACTATGGGATTAATTTCTTGGAGACAAGCGCAAAATCCAGTGCAAAATGTAGAAGAG 571
QY 525 TCATT 529
Db 572 GCATT 576

RESULT 14
US-09-834-975-885
; Sequence 885, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 885
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-885
Query Match 6.7%; Score 217.8; DB 10; Length 2497;
Best Local Similarity 65.6%; Pred. No. 1.5e-36;
Matches 318; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 45 ATGGCGAAGCAGTACGATGTGCTTCCGGCTGCTGCTGATCGGGGACTCCGGGTGGCG 104
Db 92 ATGGCGAAGCAGTACGATTTATCTTCAAGCTCTCTGATCGGGGACTCGGGGTAGGC 151
QY 105 AAGACTGCCTGCTGCGGCTTCAACGACACAGAGTTCCTCTCCGACATCTCCACC 164
Db 152 AAGACTGCCTGCTTCCGCTTCTCAGAGGAGCGCTTCAACACCACTTCATCTCCACC 211
QY 165 ATCGGTGTTGACTTTAAGATGAAGACCATAGAGGTAGCGCATCAAAAGTCGGGATACAG 224
Db 212 ATCGGAATTGATTTTAAATTTAGAACGATAGAACTAGATGGAAGAAAATTAAGCTTCAG 271
QY 225 ATCTGGGACACTGCGAGGCGAGGAGATACACAGACCATCAAAAGCAGTACTACTATCGGGCGG 284
Db 272 ATATGGGACACAGCGGGTCAGGAAGATTCCGAAACAATCACGACGCGTACTACAGAGGA 331
QY 285 GCCAGGGGATATTTTGGTCTATGACATTTAGCAGCGAGCGCTCTTACCAGCACATCATG 344
Db 332 GCCATGGGCATTTATGCTGTCTATGACATTCACAAATGAAAAATCCTTTGACAATATAAA 391
QY 345 AAGTGGGTGAGTGCAGTGCATGAGTACGACCAAGCGCTCCAGAAAGATCCTTATTGGG 404
Db 392 AATTGGATCAGAAACATTTGAAGAGCATGCTCTTCCGATGTCGAAAGAATGATCTGGGT 451
QY 405 AATAAGGCTGATGAGGACAGAAAGCGGAGGTGGGAAAGAGACAAGGCGAGCGAGTGGCG 464
Db 452 RACAATGTGATATGAATGACAAAAGACAAGTGTCAAAAGAAAGAGGAGAGCTAGCA 511
QY 465 AAGGATATGGCATGGACTTCTATGAAACAAGTGCCTGCACCACTCAACATTTAAAGAG 524
Db 512 ATTGACTATGGGATTAATTTCTTGGAGACAAGCGCAAAATCCAGTGCAAAATGTAGAAGAG 571
QY 525 TCATT 529
Db 572 GCATT 576

Db 272 ATATGGGACACACGCGGTTCAGGAAGATTCCGAACAATCACACAGCGGTACTACAGAGGA 331
QY 285 GCCCAGGGGATATTTTGGTCTATGACATTAGCAGCGGCGTCTTACCAGCACATCATG 344
Db 332 GCCATGGGCAATTATGCTGTCTATGACATCAACAATGAAAAATCCTTTCACAATATATA 391
QY 345 AAGTGGGTACGTGACGTGAGTACGCACCAAGAGCGTCCAGAAAGATCCTTATTTGGG 404
Db 392 AATTGGATCAGAAACATTGAAGAGCATGCCTCTCCGATGTCGAAAGATGATCCTGGGT 451
QY 405 AAGTGGGTACGTGACGTGAGTACGCACCAAGAGCGTCCAGAAAGATCCTTATTTGGG 404
Db 392 AATTGGATCAGAAACATTGAAGAGCATGCCTCTCCGATGTCGAAAGATGATCCTGGGT 451
QY 405 AATAAGGCTGATGAGAGCGAGAAACGGCAGGTGGGAAGAGCAAGGGCAGCAGCTGGCG 464
Db 452 AACAAATGTGATGAATGACAAAAAGACAAGTGTCAAAAAGAAAGAGGGGAGAGCTAGCA 511
QY 465 AAGGAGTATGGGATGACCTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAAGAG 524
Db 512 ATTGACTATGGGATTAAATTTCTTGAGACAGCGCAAAATCCAGTCCAAATGTAGAGAG 571
QY 525 TCATT 529
Db 572 GCATT 576

RESULT 15
US-09-834-975-894
; Sequence 894, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 894
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-894

Query Match 6.7%; Score 217.8; DB 10; Length 2497;
Best Local Similarity 65.6%; Pred. No. 1.5e-36;
Matches 318; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
QY 45 ATGGCGAAGCAGTACGATGTGCTGTTCCGGCTGCTGTCGCGGAGCTCCGGGGTGGGC 104
Db 92 ATGGCGAAGCAGTACGATGTGCTGTTCCGGCTGCTGTCGCGGAGCTCCGGGGTGGGC 151
QY 105 AAGACCTGCTGCTGTCGCCCTTCCAGGCAACGAGTTCACCTCCCTCCGACATCTCCACC 164
Db 152 AAGACCTGCTGCTGTCGCCCTTCCAGGCAACGAGTTCACCTCCCTCCGACATCTCCACC 211
QY 165 ATCGGTGTTGACTTTAAGATGAGACCATAGAGGTAGCGGCATCAAAAGTGGCGATACAG 224
Db 212 ATCGGAATTTGATTTTAAATTTAGACCATAGAGTGAAGAAATTAAGCTTCAG 271
QY 225 ATCTGGGACACTGCGAGGCGAGAGATACAGACCATCAAAAGCAGTACTATTCGGCGG 284
Db 272 ATATGGGACACAGCGGGTCCAGGAAGATTCCGAACAATCACGACAGCGTACTACAGAGGA 331
QY 285 GCCCAGGGGATATTTTGGTCTATGACATTAGCAGCGGCGCTCTTACCAGCACATCATG 344

Db 332 GCCATGGGCAATTATGCTGTCTATGACATCAACAATGAAAAATCCTTTCACAATATATA 391
QY 345 AAGTGGGTACGTGACGTGAGTACGCACCAAGAGCGTCCAGAAAGATCCTTATTTGGG 404
Db 392 AATTGGATCAGAAACATTGAAGAGCATGCCTCTCCGATGTCGAAAGATGATCCTGGGT 451
QY 405 AATAAGGCTGATGAGAGCGAGAAACGGCAGGTGGGAAGAGCAAGGGCAGCAGCTGGCG 464
Db 452 AACAAATGTGATGAATGACAAAAAGACAAGTGTCAAAAAGAAAGAGGGGAGAGCTAGCA 511
QY 465 AAGGAGTATGGGATGACCTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAAGAG 524
Db 512 ATTGACTATGGGATTAAATTTCTTGAGACAGCGCAAAATCCAGTCCAAATGTAGAGAG 571
QY 525 TCATT 529
Db 572 GCATT 576

Search completed: November 17, 2002, 20:38:39
Job time : 194 secs

07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218290.
26-JUL-2000; 2000US-0220963.
26-JUL-2000; 2000US-0220964.
14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0224521.
14-AUG-2000; 2000US-0225213.
14-AUG-2000; 2000US-0225266.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225270.
14-AUG-2000; 2000US-0225447.
14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225758.
14-AUG-2000; 2000US-0225759.
18-AUG-2000; 2000US-0226279.
22-AUG-2000; 2000US-0226681.
22-AUG-2000; 2000US-0226688.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227009.
30-AUG-2000; 2000US-0228924.
01-SEP-2000; 2000US-0229287.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.
01-SEP-2000; 2000US-0229345.
05-SEP-2000; 2000US-0229509.
05-SEP-2000; 2000US-0229513.
06-SEP-2000; 2000US-0230437.
06-SEP-2000; 2000US-0230438.
08-SEP-2000; 2000US-0231242.
08-SEP-2000; 2000US-0231243.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231413.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0232080.
08-SEP-2000; 2000US-0232081.
12-SEP-2000; 2000US-0231968.
14-SEP-2000; 2000US-0232397.
14-SEP-2000; 2000US-0232398.
14-SEP-2000; 2000US-0232399.
14-SEP-2000; 2000US-0232400.
14-SEP-2000; 2000US-0232401.
14-SEP-2000; 2000US-0233063.
14-SEP-2000; 2000US-0233064.
14-SEP-2000; 2000US-0233065.
21-SEP-2000; 2000US-0234223.
21-SEP-2000; 2000US-0234274.
25-SEP-2000; 2000US-0234997.
25-SEP-2000; 2000US-0234998.
26-SEP-2000; 2000US-0235484.
27-SEP-2000; 2000US-0235834.
27-SEP-2000; 2000US-0235836.
29-SEP-2000; 2000US-0236327.
29-SEP-2000; 2000US-0236367.
29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
02-OCT-2000; 2000US-0237040.
13-OCT-2000; 2000US-0239935.
13-OCT-2000; 2000US-0239937.
20-OCT-2000; 2000US-0240960.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241785.
20-OCT-2000; 2000US-0241786.
20-OCT-2000; 2000US-0241787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
20-OCT-2000; 2000US-0241826.
01-NOV-2000; 2000US-0244617.
08-NOV-2000; 2000US-0244674.
08-NOV-2000; 2000US-0244675.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
17-NOV-2000; 2000US-0249216.
17-NOV-2000; 2000US-0249265.
17-NOV-2000; 2000US-0249297.
17-NOV-2000; 2000US-0249299.
17-NOV-2000; 2000US-0249300.
01-DEC-2000; 2000US-0250160.
01-DEC-2000; 2000US-0250391.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251988.
05-DEC-2000; 2000US-0256719.
06-DEC-2000; 2000US-0251479.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251868.
08-DEC-2000; 2000US-0251869.
08-DEC-2000; 2000US-0251989.
08-DEC-2000; 2000US-0251990.
11-DEC-2000; 2000US-0254097.
05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and
metastasis -
Disclosure; SEQ ID NO 37097; 3071pp + Sequence Listing; English.
AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
activity, and can be used in gene therapy and vaccine production. (I)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (I) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (I) by expressing inactive proteins or to

Db 1180 GGAAGAACCATCAACCTTCTCCTGACTTACCAACCAGGAAACACGACGAGAGG 1121
QY 2453 GTGGCTCAGACTTAGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGGAGACGAGA 2512
Db 1120 GTGGCTCAGACTTAGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGGAGACGAGA 1061
QY 2513 ACTGTGAATCACTGCTTAATGAGAAAGGAGACACTTAACCTAGATGAAGCTGTGAC 2572
Db 1060 AGTGTGAATCACTGGCTAATGAGAAAGGAGACACTTAACCTAGATGAAGCTGTGAC 1001
QY 2573 TAGGCTGGAGTGTCTTCCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 2632
Db 1000 TAGGCTGGAGTGTCTTCCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATC 941
QY 2633 ACACCTGGGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGACTT 2692
Db 940 ACACCTGGGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGACTT 881
QY 2693 GAGAAGTTATATGGGCAGTGGCTCAANTCTGTGGACCAAGTATTTCAGCTTTCCTCGAAG 2752
Db 880 GAGAAGTTATATGGGCAGTGGCTCAANTCTGTGGACCAAGTATTTCAGCTTTCCTCGAAG 821
QY 2753 ATCAGCAGGGTGCCATTCACTTCTTCTCTAGCCCTTCAGGAAAGAGGACTAT 2812
Db 820 ATCAGCAGGGTGCCATTCACTTCTTCTAGCCCTTCAGGAAAGAGGACTAT 761
QY 2813 ATTGTACTGTACCCCTAGGGGTTCTGGAAGGGAAACATGGAATCAGGATTTCTATAGACT 2872
Db 760 ATTGTACTGTACCCCTAGGGGTTCTGGAAGGGAAACATGGAATCAGGATTTCTATAGACT 701
QY 2873 GATAGCCCTATCCACAAAGGCCATGACTGGGAAAGGATATGGGAGCAGAAGGAATTG 2932
Db 700 GATAGCCCTATCCACAAAGGCCATGACTGGGAAAGGATATGGGAGCAGAAGGAATTG 641
QY 2933 GGATTTTAGGTCACCTAGCTACCCCTAAACTTTTGTGGCTGCGGCATGCTTTGAG 2992
Db 640 GGATTTTAGGTCACCTAGCTACCCCTAAACTTTTGTGGCTGCGGCATGCTTTGAG 581
QY 2993 GCCAGACTGTTAAGCAGGCTGTGCTGGGCTGTTTACTCGTCACCACTCTGCACCTGCT 3052
Db 580 GCCAGACTGTTAAGCAGGCTGTGCTGGGCTGTTTACTCGTCACCACTCTGCACCTGCT 521
QY 3053 GTCTTGAGACTCAATCCAGCCAGCAGCCACCTGCTCGCTGAGCCTTCGATCTCC 3112
Db 520 GTCTTGAGACTCAATCCAGCCAGCAGCCACCTGCTCGCTGAGCCTTCGATCTCC 461
QY 3113 TGTGACGGTGAACCTGCTGTACTGTCTCGGGTCCATATATGAATTGTGACAGGGTT 3172
Db 460 TGTGACGGTGAACCTGCTGTACTGTCTCGGGTCCATATATGAATTGTGACAGGGTT 401
QY 3173 CATCTATTTTAAACACAGATGTTTACAAAATAAGATTATTTCAAACCCACC 3223
Db 400 CATCTATTTTAAACACAGATGTTTACAAAATAAGATTATTTCAAACCCACC 350

RESULT 2

AAC75813

ID AAC75813 standard; cDNA; 1666 BP.

XX AAC75813;

AC AAC75813;

XX AAC75813;

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF1368 polynucleotide sequence S8Q ID NO:2735.

DE Human; open reading frame; ORFX; detection; cytotatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressive; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX Homo sapiens.
OS WO200058473-A2.
PN 05-OCT-2000.
XX 31-MAR-2000; 2000WO-US08621.
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach M;
XX WPI; 2000-602362/57.
DR P-PSDB; AAB41604.
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders' and cardiovascular disease -
XX Claim 5; Page 1978-1979; 5507pp; English.
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressive;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus, lupus
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX SQ Sequence 1666 BP; 333 A; 529 C; 461 G; 341 T; 2 other;
Query Match 48.1%; Score 1568.2; DB 21; Length 1666;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1598; Conservative 0; Mismatches 3; Indels 18; Gaps 1;
QY 2 GCCCGCTGCCCGCCGAGTTCGCCGCGCCGCTGGCCCCCAGTATGCGAAGCAGTACGA 61
Db 48 GCCCGCTGCCCGCCGAGTTCGCCGCGCCGCTGGCCCCCAGTATGCGAAGCAGTACGA 107
QY 62 TGTGCTGTTCCGCTGCTGCTGATCGGGGACTCGGGGTGGCAAGACCTGCCCTGCTG 121
Db 108 TGTGCTGTTCCGCTGCTGCTGATCGGGGACTCCGGGGTGGCAAGACCTGCCCTGCTG 167
QY 122 CCCTCTTCCACGACAAGAGTTCACCTCTCCGACATCTCCACCATCGGTGTTGACTTTAA 181
Db 168 CCCTCTTCCACGACAAGAGTTCACCTCTCCGACATCTCCACCATCGGTGTTGACTTTAA 227

Qy	182	GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTCGGACACTGCAGG	241
Db	228	GATGAAGACCATAGAGGTAGACGGCATCAAAAGTGGGATACAGATCTCGGACACTGCAGG	287
Qy	242	GCAGGAGAGATACCACAGCATCACAAAGCAGTACTATCGGCGGCCACAGGGATATTTTT	301
Db	288	GCAGGAGAGATACCACAGCATCACAAAGCAGTACTATCGGCGGCCACAGGGATATTTTT	347
Qy	302	GGTCTATGACATTAGCAGGAGCGCTCTTTACCAGCACACATCATGAAGTGGGTCAGTGACGT	361
Db	348	GGTCTATGACATTAGCAGGAGCGCTCTTTACCAGCACACATCATGAAGTGGGTCAGTGACGT	407
Qy	362	GGATGAGTAGCCACCAAGAGGGCGTCAGAAAGATCTTATTTGGGAATAAGCGTGAAGAGA	421
Db	408	GGATGAGTAGCCACCAAGAGGGCGTCAGAAAGATCTTATTTGGGAATAAGCGTGAAGAGA	467
Qy	422	GCAGAAACGCAGGTGGGAGAGAGCAAGGGCAG-----CAGCTGGC	463
Db	468	GCAGAAACGCAGGTGGGAGAGAGCAAGGGCAGAGAAATGTCCTCTCTCAAGCTGGC	527
Qy	464	GAAGGAGTATGGCATGAGCTTCTATGAAACAAAGTGCCTGCACCAACCTCAACATTAAAGA	523
Db	528	GAAGGAGTATGGCATGAGCTTCTATGAAACAAAGTGCCTGCACCAACCTCAACATTAAAGA	587
Qy	524	GTCAATTCACGGCTGTGACAGAGCTGGTGTGCAGGCCCATAGGAAGGAGCTGGAAGGCC	583
Db	588	GTCAATTCACGGCTGTGACAGAGCTGGTGTGCAGGCCCATAGGAAGGAGCTGGAAGGCC	647
Qy	584	CCGGATCCGTGCCAGCAATGAGTTGCCACTGGCAGAGCTGGAGGAGGAGGGGCAAAACC	643
Db	648	CCGGATCCGTGCCAGCAATGAGTTGCCACTGGCAGAGCTGGAGGAGGAGGGGCAAAACC	707
Qy	644	CGAGGGCCAGCGAACTCTTTCGAAAACCTGCTGGTGCTGAGTCTGTGTGGGCAACCCCA	703
Db	708	CGAGGGCCAGCGAACTCTTTCGAAAACCTGCTGGTGCTGAGTCTGTGTGGGCAACCCCA	767
Qy	704	CACGACACCCTCTTCCCTCAGAGAGCCCGTGGGCAGACAGGGGAGCCGGGCTTTGGCC	763
Db	768	CACGACACCCTCTTCCCTCAGAGAGCCCGTGGGCAGACAGGGGAGCCGGGCTTTGGCC	827
Qy	764	TGCTGTCTCCTCTCGTGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCTCCCGCTGCC	823
Db	828	TGCTGTCTCCTCTCGTGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCTCCCGCTGCC	887
Qy	824	TGGCCCTTGAGAGGGCTCTGCTGTATCTCAAGACGCCCTCTGCCCCAGCCCGTCCACCC	883
Db	888	TGGCCCTTGAGAGGGCTCTGCTGTATCTCAAGACGCCCTCTGCCCCAGCCCGTCCACCC	947
Qy	884	TGGAGTGGTCTTCTTCAGCCTGTTTTCCCGAGCCACAGGCCTGCTACGACCCCCACGATGT	943
Db	948	TGGAGTGGTCTTCTTCAGCCTGTTTTCCCGAGCCACAGGCCTGCTACGACCCCCACGATGT	1007
Qy	944	GCGCAGAGCACTGTCTCACATCCCGCACCCACAGACAAACAGCCAGGGCTGGAGTCCAG	1003
Db	1008	GCGCAGAGCACTGTCTCACATCCCGCACCCACAGACAAACAGCCAGGGCTGGAGTCCAG	1067
Qy	1004	GCCACTTTCAGCTGCTCTTCTCCTGTCATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTC	1063
Db	1068	GCCACTTTCAGCTGCTCTTCTCCTGTCATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTC	1127
Qy	1064	CCCCACTTCTCTTCTCTGACCCCTCCCCCTCGGGTGCCTTTTCGTATCAAGACTCCTCAAA	1123
Db	1128	CCCCACTTCTCTTCTCTGACCCCTCCCCCTCGGGTGCCTTTTCGTATCAAGACTCCTCAAA	1187
Qy	1124	CCCCGTCCCCGTGTGTCTGT	1183
Db	1188	CCCCGTCCCCGTGTGTCTGT	1247
Qy	1184	CAAGGGGATGACCCAGGCTCGTGGGGAGGTTCCACCCCTTGATCCAGGAGAACCCCTCC	1243
Db	1248	CAAGGGGATGACCCAGGCTCGTGGGGAGGTTCCACCCCTTGATCCAGGAGAACCCCTCC	1307
Qy	1244	ACCTGTCTCTGGTGGGCGCAAAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCTCT	1303

Db	1308	ACCTGCTCGTGGGTGGGCCAAAGGCTACAGGGTGCCTTCTTCTCTTCCGCCACCCCA	1367
Qy	1304	CTGTCCTCATGTGCCATGGGCTGCTCCCCAGTGCACCTGCGAAAGTGGAGCATCGAGG	1363
Db	1368	CTGTCTCATGTGCCATGGGCTGCTCCCCAGTGCACCTGCGAAAGTGGAGCATCGAGG	1427
Qy	1364	TAGGAGGAAACAGCAACCGGGAGTCTCTCGAGCCTGGGGCTGCCTTACCTCTACCCATT	1423
Db	1428	TAGGAGGAAACGCAACCAAGGAGTCTCTCGAGCCTGGGGCTGCCTTACCTCTACCCATT	1487
Qy	1424	CCCGACACAGAGCTTTGCCCTTTCCTTGGCTGCCCGGCTGCCCTTTCCTTGGGGAACTGAGCTC	1483
Db	1488	CCCGACACAGAGCTTTGCCCTTTCCTTGGCTGCCCGGCTGCCCTTTCCTTGGGGAACTGAGCTC	1547
Qy	1484	AGAGGCAGGTGCTTCAGAGAAGGAAACAAATAGAGGGTGGCAGGATATAAAGTCACT	1543
Db	1548	GGAGGCAGGTGCTTCAGAGAAGGAAACAAATAGAGGGTGGCAGGATATAAAGTCACT	1607
Qy	1544	CCATTCTCTACCTCCCATGAGCATGAACACAAATTTCTCTCCACCTGGCTCCCAATTT	1602
Db	1608	CCATTCTCTACCTCCCATGAGCATGAACACAAATTTCTCTCCACCTGGCTCCCAATTT	1666
RESULT 3			
ID	ABK40069		
XX	ABK40069 standard; DNA; 7924 BP.		
AC	ABK40069;		
DT	21-MAY-2002 (first entry)		
XX	Human chemically pretreated gene sequence #76 strand 1.		
DE	Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;		
KW	cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;		
KW	UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.		
XX	Homo sapiens.		
OS	WO200202806-A2.		
PN	10-JAN-2002.		
PD			
XX			
XX	29-JUN-2001; 2001WO-EP07470.		
PF			
XX			
PR	30-JUN-2000; 2000DE-1032529.		
PR	01-SEP-2000; 2000DE-1043826.		
XX			
PA	(EPIG-) EPIGENOMICS AG.		
XX			
PI	Olek A, Piepenbrock C, Berlin K;		
XX			
DR	WPI; 2002-154757/20.		
XX			
PT	New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,		
PT	useful for detecting cytosine methylation state of genes associated		
PT	with pharmacogenomics and for therapy of diseases e.g. cancer		
XX			
FS	Claim 1; SEQ ID No 151; 24pp; English.		
XX			
CC	The invention relates to a nucleic acid comprising a sequence at		
CC	least 18 bases in length of a segment of the chemically pretreated DNA		
CC	of genes associated with pharmacogenomics according to one of the		
CC	sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B		
CC	(NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2		
CC	(NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),		
CC	MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,		
CC	NM_019899) and their complementary sequences, or a sequence (SI) chosen		
CC	from 87 sequences and their complements. The chemical pretreatment		
CC	is bisulphite treatment to convert cytosines (but not methyl-cytosines		
CC	into uracils. Also included are an oligomer (II) in particular an		
CC	oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising		

each case at least one base sequence having a length of 9 nucleotides which hybridises to or is identical to a chemically pretreated DNA of genes associated with pharmacogenomics and their complements, arranged in an array for analysing diseases associated with the methylation state (CpG) of the 87 detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The oligomers may also be used as PCR primers. The set of 87 nucleic acids and their complements is useful for diagnosis and therapy of solid tumours and cancer. The present sequence represents one the 87 DNA sequences or its complement. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

QY	1276	GGTGCTTCTCTCTCCGCCACCCCTGCTCCCTCATGTGCCATGGGCTGCCCTCCCC	1335
Db	1	GGTGTTTTTTTTTTTTTTTATTTTATTTATTTGTTTTATGCTGTATGGGTTGTTTTTT	60
QY	1336	AGTGACCTCCGAAGTGGAGCATCAGAGTAGGAGGAAACACCAACCCGGGAGTCCCTGA	1395
Db	61	AGTGATTTCCGAAGTGGAGTATCAGAGTAGGAGGAAACCGTAATTAGGAGTTTCGA	120
QY	1396	GCCTGGGGTGCCTACCTACCTACCCATCCCGACAGAGCTTTGCCCTTGTCTGCTCC	1455
Db	121	GTTTGGGGTGTGTTTTATTTTATTTTATTTTATTTTATTTTATTTTGTGTTGTTT	180
QY	1456	CGCCCTCGCTCTTTGGGGAACCTGAGCTCAGAGGAGCTGCTTCAGAGAAAGAAACAAAT	1515
Db	181	TCGTTTGTGTTTTTTGGGGAATGAGTTTATAGAGTGTGTTTATAGAGAAATTAAT	240
QY	1516	GAGGGTGGCAGGATPAAAGTACCTCCATCTCTACCTCCCATGCCAGCATPAGAACACA	1575
Db	241	GAGGGTGGTAGGATPAAAGATTTATTTTATTTTATTTTATTTTATTTATGATGATATA	300
QY	1576	ATTTCTCTCCACTGCTCCCAATTTAAAGATGTGGCAAGGCTGTGGGTACTCCAG	1635
Db	301	ATTTTTTTTTTTGTTGTTTTTAAATTTAAAGATGTGGATTAAGGTTTGGGTATTTTAG	360
QY	1636	GGCAAGGAGAGCCCTGGGTCAGTGACACTGTTCAGCCCAACCATCCTCCACAAAGG	1695
Db	361	GGTAAGGAGAGTTTGGGTTAGTCATATTGTTAGGTTAAATATTTATTTATAAAGG	420
QY	1696	GACATTTGGAATGAAGACTAGTCCCTATGTATCAGGTTAAGACAAGGAGAGCTGG	1755
Db	421	GAGTATTTGGAATGAAGGATTTAGTTTTTATGTTATGTTAAGAGTAAGGAGAGTTGG	480
QY	1756	CGAGGACAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCAGGCTCCTAGGCC	1815
Db	481	TTAGGATAGTAGTTGTTATAGTAGAGGGGAATGTAGTAATAGTAGGTTTTTAGGTTT	540
QY	1816	CATCTTCCATTTCTTAGTGAAGAGCAATTTCCCTCAGACTCCAGCGGAGGACTGAGC	1875
Db	541	TATTTTTTATTTTATAGTGAAGAGATTTTTTTTTTAGATTTTTTAGCGGAGGATTGAGT	600
QY	1876	CTAGCCTTCAGCAACCAAGTTCTCTCGGACCCCAAGTTTATGGAGAGGCAAGAC	1935
Db	601	TTAGTTTTTATTTTATAGTAAAGTTTTTTTTTGGGATTTAAAGTTTTATGGAGAAGGTAAGAT	660
QY	1936	TTTCATGGAGAGAGAGGAGCCCTGGGTAGAAAGCTGTGCTGTCTCTCTTTGGCC	1995
Db	661	TTTATGGGAAGAGAGAGGAGGTTTTTGGGTAGAAAGCTTTTGGTGTGTTTTTTTGGTT	720
QY	1996	TTTAAAGCAAGCGCTCATCTTCCCTCTACCTCTAGAGGCTTGAAGGTTTGGCCAAAC	2055
Db	721	TTTAAAGATAAGCGTTTATTTTCTTTTATTTTATTTTATTTTATTTGATAGGTTTTCAGG	780
QY	2056	ACACTGTGGCTACAGGTGGAGGGAAGAGGACTCCTCTCCAGAGTGTATGTTCCAGAA	2115

Db	781	ATATTTGCTGTTATAGCTGAGGGAAGAGGATTTTTTTTTTTTATAGTGTATGTTTAGGAA	840
QY	2116	GTTTCTTTAAACCCCAATGATGGCCCAAGAGTAGCTCTAGAGGCCCTTTAAAGACGGAACA	2175
Db	841	GTTTTTTTTTAAATTTTATATGTTTAAAGAGTAGTTCGTAGAGGTTTTTTTAAAGACGGAATA	900
QY	2176	AGTAATTTTACCAGTTCTACTGGGGTTCTGCCACCCGTCCTCCCAAGGTGGCGAGGCTAGG	2235
Db	901	AGTAATTTTATAGTTTATTTGGGGTTTTTGTATTCGTTTTTAAGGTGGCGAGGTTAGG	960
QY	2236	AAGAGGTCATTTTAAAGCCACACATTTAGCTGCACCTGGTGGCTGCAGCCCAAAACAAGA	2295
Db	961	AAAGAGGTTATTTTAAAGTTATATATTTAGTTGCTATTTGCTGTTGTTTAAATAAAGA	1020
QY	2296	ACTGGTGTGTAGTATTCATCAACTAAAGAACCAAAATCCAGGCCACTCATATGTGAAGA	2355
Db	1021	ATTGGGTGTGTAGTATTTATTTAAATTAAGNATTTAAATTTAGGGTATTTATATGTGAGGA	1080
QY	2356	TAAGAACCTTACTTCTTACTCTCCAAAAGAGTGGGGAAAGAACCACTCAACCTTTC	2415
Db	1081	TAAGAAATTTTATTTTATTTTATTTTAAAGAGTGGGGAAGAATTTATTAATTTTTT	1140
QY	2416	CTCTGACTTACCAACACAGGAAACACAGGAGAGGTTGGCTCAGGACTTAGGNCAGG	2475
Db	1141	TTTTTTCATTTTAAATTTAGGAAATATAGTAGAGAGGTTGGTTTAGGATTTAGGGATAGG	1200
QY	2476	GTATAGCTTAGATGTTGGAAGCAAGAGAGCAGCAAGCTTGTAAATCACTGCTCAATGA	2535
Db	1201	GTATAGTTTATAGTGTGGAAAGTAAAGAGAGTAGAGAACTTCTAAATTTATGTTAATCA	1260
QY	2536	GAAGAGGAGACAGCTAATCTTAGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCCCTGAA	2595
Db	1261	GAAGAGGAGATAGTTAAATTTTAGGATGAAGTTGCTGATAGTTGGAGTTGTTTTTGA	1320
QY	2596	GATGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGGCTAGGGAAGTAGCT	2655
Db	1321	GATGGGATTTTGGGTATTAAGATTTATGTTATATTTATTTGGGTTAGGGAAGTAGGT	1380
QY	2656	GATCCAGCCCTCAAGCTGTCTCTCAGCCAGGAGCTTGAGAAGTTATATTTGGSCAGTGC	2715
Db	1381	GATGTTAGTTTAACTGTTTGTAGTTAGGATTTAGAGAGTTATATTTGGGTAGTGT	1440
QY	2716	TCAATCTCTGGACCAAGTATTTTCAGCTTTCCCTGGAAGATCAGCAGGTTGCCATTCATTTG	2775
Db	1441	TTTAATTTCTGGATTTAGTATTTTATTTTGGGTTAGGATTTAGGTTAGGTTATTTATTG	1500
QY	2776	TCTTCTCTCTAGCCCTCAGGAAAGAGGACTATATTTGTACTGTACTCCCTAGGGGTT	2835
Db	1501	TTTTTTTTTTTTTATGTTTTTTTAGGAAAGAGGATTTATTTGTTATTTTAGGGTT	1560
QY	2836	CTGGAAGGAAACATGGAATCAGGATTTCTATAGACTGTATAGGCCCTATCCCAAGGCC	2895
Db	1561	TTGGAAGGAAATATGGAATTTAGGATTTATAGATTTAGGTTTATTTATTAAGGGTT	1620
QY	2896	ATGACTGGGAAAGGATTTGGGACCAAGAGAAATTTGGGATTTTAGGTCGACTAGCT	2955
Db	1621	ATGATTTGGGAAAGGATTTGGGATTTAGAGGAAATTTGGGATTTTAGGGTGTAGTACGT	1680
QY	2956	CACCTTAAACTTTTGTGGCTGGGCTGCTCTTGGAGGCCACAGCTGTTAAAGCAGGCTCT	3015
Db	1681	TATTTTAAATTTTGGTGGTTGGGATGTTGTTGAGTTTAGATTTGTTAAATTTAGGTTT	1740
QY	3016	GCTGGCTGTTTACTCGTCACCACTCTGCACCTCTGCACCTGCTGTCTTGAGACTCCATCCAGCC	3075
Db	1741	GTTGGTTTGTATTATTCGTTATTTATTTTGTGTTGTTTGTGTTTATTTATTTAGTTT	1800
QY	3076	AGCAGCCACCTGCTGCTGAGCCTCCATATCTCCCTGTGAGCGGTGACTTCGTGTAC	3135
Db	1801	AGTAGCTATTTTGTGTTTTTGTAGTTTTTATTTATTTTGTGACGGGTGAATTTTCGTGTAT	1860
QY	3136	TGTGCTCTCGGTCATATATGAATTTGTGAGCAGGTTTCATCTATTTTAAACACAGATGTT	3195

[illegible]

Db 7562 ACAAAAAAAAAAACCCTAAATCAATACACTATCAAAACCAACCATACACTCCACAAAAAAA 7503
QY 1698 GCATTGGAAATGAAGGACTAGCTCCTATGTATCAGGTTTAAGACGAAGGAGCTGGCC 1757
Db 7502 ACATTTAAAAATAAAAAAATACTACTCCTATATATCAATTAATAAACAACAAAAAATACTAAC 7443
QY 1758 AGGACAGAGTTTGCACAGAGAGGGAATGATGACAGCAGGCGCTCCTAGGCCCA 1817
Db 7442 AAAAAACAATTTACACAACAAAAAATAATACACAACAAAAACCTCCTTAAACCCCA 7383
QY 1818 TCTTCCATTTCTAGTAAAGAGACATTTCTCAGACTCCAGGCGGAGGACTGAGCCT 1877
Db 7382 TCTTCCATTTCTAAATAAAAAAATACTTCTCAAACTCCCAACGAAAACTAAACCT 7323
QY 1878 AGCCTTCAGCAACCAAGGTTCTCCCTGGGACCCAAAGTTTATGGGAGAGGGCAAGACTT 1937
Db 7322 AACCTTCACACCAAAATTTCTCTTAACCCCAAAATTTATAAAAAACAACAAAACTT 7263
QY 1938 CATGGAGAGAGAGAGGAGGCGCTGGGTAGAAACGCTTGTGCTGTTCTCTTTGGCCCTT 1997
Db 7262 CATAAAAAAATAAAAAAACCCTTAATAAATAAAGCTTAATACTATTCTCTTTAACCTT 7203
QY 1998 TAAGCAAAAGCCTCATCTTGGCCCTCTACCTCCTGATAGGCTTGAGGCTTGGCCAAACAC 2057
Db 7202 TAAACAAACACGCTCATCTTACCCCTCTACCTCTAATAAATTAATAAATTTTACCAACCCAC 7143
QY 2058 ACTGTGGCTACAGGTGGAGGAGGAGGACTCCTCTCCAGAGTGTATGTTTCAGGAAGT 2117
Db 7142 ACTATACTACAATAAATAAAAAAANAACCTCCTCTCCAAATACTATATTCAAAAAT 7083
QY 2118 TTTCTTAACCCCATATGGCCCAAGAGTAGCTCGTAGGAGGCGCTTTAAAGACGGAACAG 2177
Db 7082 TTTCTTAACCCCATATAACCCCAAAATAACTCGTAAAAAACCTTTAAAAACGGAACAAA 7023
QY 2178 TAATTTACAGTTCTACTGGGTTCTGCCCACCGTCCCAAGGTGGCGGAGGCTTAGGAA 2237
Db 7022 TAATTTACCAATTTACTTAAATTTCTACCCACCGTCCCAAAATAAAGAAACCTTAAAAA 6963
QY 2238 GAGGCTCATTTTAAGCCACACATTAAGCTGCACTGGCTGCGAGCCCAACCAACAAAGAAC 2297
Db 6962 AAAAAATCATTTTAACACACATTAACCTACACTAGCTAGCTACCAACCAACCAACAAAC 6903
QY 2298 TGGGTGTTGAGTATCACTAAGAACCAAAATCCAGGCGCACTATATGTGAAGGATA 2357
Db 6902 TAAATATTAAATTTATCACTAATAAACCACCAAAATCCAAACACTATATAAAAAATA 6843
QY 2358 AGAACCTCACTTCTTACTCTCCAAAGAGAGTGGGGAAGCAACCAACCACTTCTCT 2417
Db 6842 AAAAACTCACTTCTTACTCTCCAAAGAGAGTGGGGAAGCAACCAACCACTTCTCT 6783
QY 2418 CCTGACTTACCAAAACGAGGAACAGCAGGAGGCTGCTCAGGACTTAGGACAGGCT 2477
Db 6782 CTTAATTTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 6723
QY 2478 ATAGCTTAGATGGTGGAAAGCAAGAGAGAGGAGGAGTGTAAATCACTGGCTAATGAGA 2537
Db 6722 ATAACCTTAAATTAATAAAAAACCAAAAAACCAAAAAATTAATAATCACTAATAAAAA 6663
QY 2538 AAAGGACAGACTAATCTAGGATGAGCTGACTAGGCTGGAGTGGCTTCTTGAAGA 2597
Db 6662 AAAAAACCAACTAATCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6603
QY 2598 TGGGACTCCTTGGGTATCAAGACCTATGCCACATCACACTGGGCTAGGGAAGTAGGTA 2657
Db 6602 TAAACTCCTTAAATATCAAAACCTTATACCATACACTAAACCTAATAAATAAATAAATAA 6543
QY 2658 TGGCAGCCCTCAAGTGTCTTTCAGCCAGGAGTGTGAGAAGTTATATTGGGAGGTGGCTC 2717
Db 6542 TACCAACCCCTCAATCTATCTTCAACCAAAAACTTAAAAAATTAATTAACCAATAACTC 6483
QY 2718 CAATCTGTGGACAGTATTTTCACTTCTCCCTGAAAGATCAGGCGGCTCATTTGTC 2777
Db 6482 CAATCTATAACCAATATTTTCACTTCTCCCTAAAAATCAACCAAAAAATACCATTCATTATC 6423

QY 2778 TTTCTCTCTAGCCCTCCAGGAAAGAGGACTATATTTGTACTGTACCTTAGGGTTCT 2837
Db 6422 TTTCTCTCTACCCCTCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6363
QY 2838 GGAAGGGAACATGGAATCAGGATTTCTATAGCTGATAGGCGCTTATCACAAGGGCCAT 2897
Db 6362 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6303
QY 2898 GACTGGAAAAAGTATGGGAGCAGAGGAATTTGGGATTTTAGGTGTCAGCTACGCTCA 2957
Db 6302 AACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6243
QY 2958 CCTTAACTTTTGGTGGCTGGGCGATGCTTTGAGGCCAGGACTGTGTAAGCAGGCTCTGC 3017
Db 6242 CCTTAACTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 6183
QY 3018 TGGCTGTTTACTCTGACCACTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3077
Db 6182 TAACTTATTTACTCTGACCACTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6123
QY 3078 GCACGCCACTCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3137
Db 6122 ACACGCCACTCTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 6063
QY 3138 TGTCTCGGTCCATATATGATTTGTGAGCAGGCTTCACTATTTTAAACACAGATGTTTA 3197
Db 6062 TATCTCGAATCCATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6003
QY 3198 CAAATAAAGATTTTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3257
Db 6002 CAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5943

RESULT 6

ABL34139/c
ID ABL34139 standard; DNA; 7924 BP.
AC ABL34139;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 2112.
DE Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytosinatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX
XX 30-JUN-2000; 2000DE-1032529.
PR
XX
XX 01-SEP-2000; 2000DE-1043826.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2002-130909/17.
DR
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -

xx
ps Claim 1; SEQ ID NO 2112; 32pp + Sequence Listing; German.
xx
cc The present invention provides a number of human immune system associated
cc genes which are modified by the methylation of cytosines. The sequences
cc can be used in the diagnosis and treatment of immune system disorders,
cc including eye diseases such as retinopathy, neovascular glaucoma and
cc macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
cc leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
cc rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
cc diseases. The present sequence is a gene of the invention.
xx
sq Sequence 7924 BP; 1973 A; 83 C; 1827 G; 4041 T; 0 other;

Query Match 35.6%; Score 1160.8; DB 24; Length 7924;
Best Local Similarity 74.1%; Pred. No. 2.9e-271; Indels 0; Gaps 0;
Matches 1468; Conservative 0; Mismatches 512;

QY 1278 TGCTTCTCTCTCCCTCCACCCCTGCTCCCTATGTCCTATGGCCCTGCTCCCTCCAG 1337
DB 7922 TACTTCTCTCTCCCTCCACCCCTGCTCCCTATGTCCTATGTCCTATGTCCTCCCA 7863

QY 1338 TGACCTCGGAAGTGGAGATCGAGTAGGAGGAACAGCAACCGGGAGTCTCGAGC 1397
DB 7862 TAACCTACGAAATAAACAATCGAAATAAATAAATAAATAAATAAATAAATAAATAA 7803

QY 1398 CTGGGGCTGCCCTACCTCTACCCATTCGCCGACGAGCTTTCGCCCTTGCTTGCTGCC 1457
DB 7802 CTAAACTACCTTACCTCTACCCATTCGCCGACGAGCTTTCGCCCTTGCTTGCTGCC 7743

QY 1458 GCCTGCTCTTTGGGGAAGTGTGCTCAGAGGAGGCTGTCTCAGAGAAGGAACAAATGA 1517
DB 7742 GCCTACCTCTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7683

QY 1518 GGGGTGCGAGGATAAAGTACCTACCTCTCTACCTCCATGTCAGATGAACACAT 1577
DB 7682 AAAATAACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7623

QY 1578 TTCTCTCCACTGCTCCCAATTTAAAGATGTGGACCAAGGCTGTGGGTACTCCAGGG 1637
DB 7622 TTCTCTCCACTGCTCCCAATTTAAAGATGTGGACCAAGGCTGTGGGTACTCCAGGG 7563

QY 1638 GCAAGAGAGCCCTGGGTCAGTGCACATCTCAGGCCAACCATGCATCCACAAAGGGA 1697
DB 7562 ACAAAAAAACCCCTAAATCAATAACACTATCAAAACCAACATACACTCCACAAAAA 7503

QY 1698 GCATTTGGAATGAGGACTAGCTCTCTATCATCAGTTAAGCAAGGAGAGCTGCC 1757
DB 7502 ACATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7443

QY 1758 AGGACAGCAGTTTGCACAGCAGAGGGAATGTAGCAACAGAGGCTCTCTAGGCCCA 1817
DB 7442 AAAACAACAATTACACACAAAAATAAATAAATAAATAAATAAATAAATAAATAA 7383

QY 1818 TCTTCCATTTCTAGGTAAGAAGCATTTCTCAGACTCCAGGCGGAGGACTGAGCCT 1877
DB 7382 TCTTCCATTTCTAGGTAAGAAGCATTTCTCAGACTCCAGGCGGAGGACTGAGCCT 7323

QY 1878 AGCTTCAGCAACCAAGTTCTCTGGGACCAAGTTTATGGAGAGGCAAGACTT 1937
DB 7322 AACCTTCACACCAAAATTTCTCTAAACCCCAAAATTTATATAAATAAATAAATAA 7263

QY 1938 CATGGGAAGAGAGGAAGGCTGCTGAGAACGCTGCTGCTCTCTTTGGCCTT 1997
DB 7262 CATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7203

QY 1998 TAAGACAAGCGCTCATCTTTCCTCTACCTCTGATAGGCTTGGAGGTTGGCAACAC 2057
DB 7202 TAAACAACAAAGCTCATCTTACCTCTACCTCTGATAGGCTTGGAGGTTGGCAACAC 7143

QY 2058 ACTGTGCTACAGTGGAGGAGAGACATCTCTCCAGAGCTGCTCTCAGGAGT 2117
DB 7142 ACTATAACTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7083

QY 2118 TTCTTTAACCCCATATGCCCCAAGAGTAGCTGTAGGAGGCCCTTTTAAAGACGAACAAG 2177
DB 7082 TTCTTTAACCCCATATATACCCCAAAATAACTCGTAAAAACCCCTTTTAAACAGAAACAA 7023

QY 2178 TAATTTACAGTTCTACTGGGTTCCTGCCACCGCTCCCAAGGTGGCGAGGCTTAGGA 2237
DB 7022 TAATTTACCAATTTCTACTAAAAATTCCTAGCCCGTCCCAAAATAAACAAGAACTTAA 6963

QY 2238 GAGGGTCATTTTAAGCCACACATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2297
DB 6962 AAAAATCATTTCTTAAACACACATTAACCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 6903

QY 2298 TGGGTCTTGAGTATTCATCAACTTAAGAACAACCAAAATCCAGGGCACTCATATGTGAAGGATA 2357
DB 6902 TAATATTAATATTCATCAACTTAAGAACAACCAAAATCCAAACACTCATATATAAATAA 6843

QY 2358 AGAACCTACTTCTCTTACTCTCCCAAAAGAGTGGGGAAGAACCATCAAACTTTCCT 2417
DB 6842 AAAACCTCTCTTCTTACTCTCCAAATAAATAAATAAATAAATAAATAAATAAATAA 6783

QY 2418 CCGTACTTACCAACACAGGAACACAGAGAGGCTGCTCAGGACTTAGGACAGGCT 2477
DB 6782 CCTACTTACCAACCAACCAAAACCAAAATAAATAAATAAATAAATAAATAAATAA 6723

QY 2478 ATAGCTTAGATGGTGAAGCAAGGAGAGCAGGAAGTTGTAAATCACTTGGCTTAATGAGA 2537
DB 6722 ATAACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6663

QY 2538 AAGGAGACAGCTAATCTTAGGATGAAGCTGTGCTAGGCTGAGGCTTCTCTTGAAGA 2597
DB 6662 AAAAAAACCACTTAACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6603

QY 2598 TGGGACTCTTGGGTATCAAGACTTATGCACATACACTGGGGCTAGGGAAGCTAGGTGA 2657
DB 6602 TAAACTCTCTTAAATCAAAACCTATACCACTACACTTAAACCTTAAATAAATAAATAA 6543

QY 2658 TGCCAGCCCTCAAGTCTGTCTTCAGCCAGGACTTGAGAAAGTTATATTTGGGCACTGGCT 2717
DB 6542 TACCAACCCCTCAAACTATCTTCAACCAAAACCTTAAATAAATAAATAAATAAATAA 6483

QY 2718 CAATCTGTGACACAGTATTTCACTTTCCCTGAAGATCAGGCGAGGCTGCATTCATGTC 2777
DB 6482 CAATCTATAACCCATATTTCACTTTCCCTTAAATAAATAAATAAATAAATAAATAA 6423

QY 2778 TTTCTCTCTAGCCCTCAGGAAAGAGACTATATTTCTACTGTACCTAGGCTTCT 2837
DB 6422 TTTCTCTCTAGCCCTCAAAACCAAAACCTATATTTATATATACCTTAAATAAATAA 6363

QY 2838 GGAAGGAAACATGGAATCAGGATTCATAGACTGATAGGCCCTATCCACAGGCGCAT 2897
DB 6362 AAAAAAACCACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6303

QY 2898 GACTGGGAAGGATGAGGAGCAGAGGAGAAATTTGGGTTAGGCTGACGCTACGCTCA 2957
DB 6302 AACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6243

QY 2958 CCCTAAACTTTTGGTGGCCCTGGGCGCATGTCTTGGAGGCCAGACTGTTTAAAGCAGGCTCTGC 3017
DB 6242 CCCTAAACTTTTAAATTAACCTTAAACATATCTTAAACCCCAAACTATTAACCAAACTCTAC 6183

QY 3018 TGCCCTGTTTACGTGACCACTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3077
DB 6182 TAACCTATTTACTGTCACCACTCTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6123

QY 3078 GCAGGCCACTGCTCTGAGCCCTCCACTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3137
DB 6122 ACAGGCCACTGCTCTTAAACCTCCACTATCTCTCTTAAACCTTCTTCTTCTTCTTCTTCTTCT 6063

QY 3138 TGCTCTGGGTCCCATATATGAATTTGTGAGCAGGCTTCTATCTTATTTAAACACAGATGTTTA 3197
DB 6062 TATCTGAATCCATATATTAATTAACAAATTCATCTATTTTAAACACAAATATTTA 6003

RESULT 8
AAZ24400
ID AAZ24400 standard; cDNA; 895 BP.
XX
AC AAZ24400;
XX
DT 14-FEB-2000 (first entry)
XX
DE Human bladder tumour cDNA library derived EST 12.
XX
KW Expressed sequence tag; human; bladder; tumour; cancer; cytostatic;
KW treatment; gene therapy; EST; ss.
XX
OS Homo sapiens.
XX
PN DE19818619-A1.
XX
PD 28-OCT-1999.
XX
PF 21-APR-1998; 98DE-1018619.
XX
PR 21-APR-1998; 98DE-1018619.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E;
XX WPI; 1999-612028/53.
XX
DR New nucleic acid sequences expressed in bladder tumor tissue, and
XX derived polypeptides, for treatment of bladder tumor and identification
XX of therapeutic agents -
XX
PS Claim 3; Page 69; 132pp; German.
XX
CC This invention describes novel polypeptide fragments (I) and the
CC polynucleotides (II) that encode them that are highly expressed in a
CC human bladder tumour and which have cytostatic activity. (II) are used
CC for recombinant expression of (I) and to isolate complete genes. (I) are
CC used to identify agents suitable for treatment of bladder cancer, to
CC directly treat this form of cancer (including expression from gene
CC therapy vectors) or are used in a preparation for cancer treatment. (I)
CC is also used for the generation of specific antibodies. (II) are
CC identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, and
CC therefore reduces the number of failures associated with the fact that
CC ESTs from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AAZ43260-243309 represent expressed sequence tag (EST)
CC fragments isolated from a human bladder tumour cDNA library which encode
CC the proteins represented in AAY66143-Y66198.
XX
SQ Sequence 895 BP; 249 A; 205 C; 228 G; 213 T; 0 other;

Query Match 26.6%; Score 866.2; DB 20; Length 895;
Best Local Similarity 99.5%; Pred. No. 3.8e-200;
Matches 879; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2358 AGAACCCTACTTCCTTACTCTCCAAAAGAGTGGGGAAGAACCATCAACCTTCTCT 2417
DB 14 AGAACCCTACTTCCTTACTCTCCAAAAGAGTGGGGAAGAACCATCAACCTTCTCT 73
QY 2418 CTGACTTACCAACAGCAAGAACAGCAGGAGGGTGCTCAGGACITAGGACAGGT 2477
DB 74 CTGACTTACCAACAGCAAGAACAGCAGGAGGGTGCTCAGGACITAGGACAGGT 133
QY 2478 ATAGCTTAGTGGTGAAGCAAGAGAGAGGAGTGTAAATCACTGGCTAATGAGA 2537
DB 134 ATAGCTTAGTGGTGAAGCAAGAGAGAGGAGTGTAAATCACTGGCTAATGAGA 193
QY 2538 AAAGGAGACAGCTAACTCTTAGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTTGAAGA 2597

DB 194 AAAGGAGACAGCTAACTCTAGGATGAAGCTGTGACTAGCTGAGTGTCTCTTGAAGA 253
QY 2598 TGGGACTCCTTGGGTATCAAGACCTATGCCATCATCACTGGGGCTAGGAAGTAGGTGA 2657
DB 254 TGGGACTCCTTGGGTATCAAGACCTATGCCATCATCACTGGGGCTAGGAAGTAGGTGA 313
QY 2658 TGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGAGCTTGAGAAGTTATATTGGGAGTGGCTC 2717
DB 314 TGCCAGCCCTCAAGTCTGTCTTTCAGCCAGGAGCTTGAGAAGTTATATTGGGAGTGGCTC 373
QY 2718 CAATCTGTGGACCAAGTATTTTCAAGCTTTCCTTGAAGTATCAAGGAGGTGCCATTCATGTC 2777
DB 374 CAATCTGTGGACCAAGTATTTTCAAGCTTTCCTTGAAGTATCAAGGAGGTGCCATTCATGTC 433
QY 2778 TTTCTCTCTAGCCCTCAGGAAAGAGGAGCTATATTTGTACTGTACCTAGGGGTTCT 2837
DB 434 TTTCTCTCTAGCCCTCAGGAAAGAGGAGCTATATTTGTACTGTACCTAGGGGTTCT 493
QY 2838 GGAAGGAAACATGGAATCTATAGACTGATAGGCCCTATCCACAGGGCCAT 2897
DB 494 GGAAGGAAACATGGAATCTATAGACTGATAGGCCCTATCCACAGGGCCAT 553
QY 2898 GACTGGGAAAGGTATGGGAGCAGAGAGAAATTTGGGATTTTAGGTGAGCTACGCTCA 2957
DB 554 GACTGGGAAAGGTATGGGAGCAGAGAGAAATTTGGGATTTTAGGTGAGCTACGCTCA 612
QY 2958 CCTTAACTTTTGGTGGCTGGGCGATGCTTGAGGCCAGAGCTGTTAAGCAGGCTCTGC 3017
DB 613 CCTTAACTTTTGGTGGCTGGGCGATGCTTGAGGCCAGAGCTGTTAAGCAGGCTCTGC 672
QY 3018 TGGCTGTCTTACTCGTCACACCTCTGCACCTGCTCTTGTGAGACTTCCAGCCGCCAG 3077
DB 673 TGGCTGTCTTACTCGTCACACCTCTGCACCTGCTCTTGTGAGACTTCCAGCCGCCAG 732
QY 3078 GCAGCCACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGGTGAATTCGTGTACTG 3137
DB 733 GCAGCCACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGGTGAATTCGTGTACTG 792
QY 3138 TGTCTCGGTCCATATATCAATTTGTGAGCAGGTTTCATCTATTTTAAACACAGATCTTTA 3197
DB 793 TGTCTCGGTCCATATATCAATTTGTGAGCAGGTTTCATCTATTTTAAACACAGATCTTTA 852
QY 3198 CAAAATAAGATTATTTCAACACACCAAAAAAAGAAAAA 3240
DB 853 CAAAATAAGATTATTTCAACACCAAAAAAAGAAAAA 895
RESULT 9
AAZ27472
ID AAZ27472 standard; cDNA; 566 BP.
XX
AC AAZ27472;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 507.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
XX acquired immune deficiency syndrome.
OS Homo sapiens.
XX
PN WO200154733-A1.
XX

PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01312.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-02114886.
PR 30-JUN-2000; 2000US-02115135.
PR 07-JUL-2000; 2000US-02116647.
PR 07-JUL-2000; 2000US-02116880.
PR 11-JUL-2000; 2000US-02117487.
PR 11-JUL-2000; 2000US-02117496.
PR 14-JUL-2000; 2000US-02118290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
XX
PI
XX
DR

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465460/50.

DR P-PSDB; AAU17555.
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders
XX
XX
XX Claim 1; SEQ ID No 507; 880pp; English.
PS
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Query Match 16.4%; Score 534.2; DB 22; Length 566;
Best Local Similarity 98.9%; Pred. No. 1.3e-119;
Matches 533; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 GCCCGCTCCGCCCGCGAGTTCCTCCGCCCGCGCTGCGCCCGAGTCATGGCGAGCAGTACGA 61
Db 28 GCCCGCTCCGCCCGCGAGTTCCTCCGCCCGCGCTGCGCCCGAGTCATGGCGAGCAGTACGA 87
QY 62 TGTGCTGTTCGGCTGCTGATCGGGGACTCCGGGTGGCGAGACCTGCTGCTGTG 121
Db 88 TGTGCTGTTCGGCTGCTGATCGGGGACTCCGGGTGGCGAGACCTGCTGCTGTG 147
QY 122 CGCCTTCACCGACACGAGTTCCTCTCGCACATCTCCACATCGGTGTGTGACTTAA 181
Db 148 CGCCTTCACCGACACGAGTTCCTCTCGCACATCTCCACATCGGTGTGTGACTTAA 207
QY 182 GATGAAGACCATAGAGGTAGAGCGCATCAAGTGGCGGATACAGATCTGGGACATGCGAGG 241
Db 208 GATGAAGACCATAGAGGTAGAGCGCATCAAGTGGCGGATACAGATCTGGGACATGCGAGG 267
QY 242 GCAGGAGATACACAGACCATCACAAAGCAGTACTATCGGGGGCCCGAGGGATATTTT 301
Db 268 GCAGGAGATACACAGACCATCACAAAGCAGTACTATCGGGGGCCCGAGGGATATTTT 327
QY 302 GGTCTATGACATTAGCAGCGCGCTCTTACCAGCACATCATGAAGTGGTTCAGTGACGT 361
Db 328 GGTCTATGACATTAGCAGCGCGCTCTTACCAGCACATCATGAAGTGGTTCAGTGACGT 387
QY 362 GGATGAGTACGCACAGAGCGCTCCAGAAATCTTATTTGGGAATAAGGCTGATGAGGA 421
Db 388 GGATGAGTACGCACAGAGCGCTCCAGAAATCTTATTTGGGAATAAGGCTGATGAGGA 447
QY 422 GCAGAAACGGCAGGTGGGAAGAGACAAAGGCGAGCGCTGGCGAAGAGATATGGCATGA 481
Db 448 GCAGAAACGGCAGGTGGGAAGAGACAAAGGCGAGCGCTGGCGAAGAGATATGGCATGA 507
QY 482 CTTCTATGAACAAGTGGCTGCACCAAGCTCAACATTAAGAGTCAATTCACCGCTCTGA 540
Db 508 CTTCTATGAACAAGTGGCTGCACCAAGCTCAACATTAAGAGTCAATTCACCGCTCTGA 566

RESULT 10
ABK44752
ID ABK44752 standard; cDNA; 481 BP.
XX
XX AC ABK44752;
XX
XX DT 05-JUN-2002 (first entry)
XX
XX cDNA encoding colon tumour protein, SEQ ID No 303.
DE
DE Human; colon tumour; vaccine; colon cancer; immunogenic;
KW immunotherapy; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200212328-A2.
PN
PD 14-FEB-2002.
XX
XX 31-JUL-2001; 2001WO-US24218.
PF
PR 03-AUG-2000; 2000US-223283P.
PR 28-MAR-2001; 2001US-279763P.
PR 29-JUN-2001; 2001US-302051P.
XX
XX (CORI-) CORIXA CORP.
PA
PI King GE, Meagher MJ, Xu J, Secríst H;
XX
XX WPI; 2002-241739/29.
DR
XX New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT for diagnosing, preventing, and treating colon cancer, and as markers
PT for the progression of cancer
XX
PS Claim 1; SEQ ID No 303; 147pp; English.
XX
XX The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
CC compositions, such as vaccines, for the diagnosis, prevention, and
CC treatment of colon cancer. Polynucleotide sequences may be used as
CC hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. The compositions are useful for stimulating an
CC immune response against cancer, particularly for the immunotherapy of
CC colon cancer, and as markers for the progression of cancer.
CC ABK44450-ABK46237 represent coding sequences of human colon tumour
CC proteins of the invention.
CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
CC for this patent did not form part of the printed specification but was
CC supplied by the European Patent Office.
XX
XX Sequence 481 BP; 111 A; 113 C; 130 G; 125 T; 2 other;
Query Match 14.7%; Score 477.4; DB 24; Length 481;
Best Local Similarity 99.4%; Pred. No. 7.3e-106;
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2582 GTTGCTTCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCACATCACACTGGGG 2641
Db 1 GTTGCTTCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCACATCACACTGGGG 60
QY 2642 CTAGGGAAGTAGTGTATGCCAGCCCTCAAGTCTGCTTCAGCCAGGACTTCAGAGTTA 2701
Db 61 CTAGGGAAGTAGTGTATGCCAGCCCTCAAGTCTGCTTCAGCCAGGACTTCAGAGTTA 120
QY 2702 TATTGGGAGTGGCTCCCAATCTGTGGACACAGTATTTCAGCTTTCCCTGAAGATCAGGCAG 2761
Db 121 TATTGGGAGTGGCTCCCAATCTGTGGACACAGTATTTCAGCTTTCCCTGAAGATCAGGCAG 180
QY 2762 GGTGCAATTCATTGTCTTTCTCTCTAGCCCCCTCAGGAAAGAGACTATATTTGTACT 2821
Db 181 GGTGCAATTCATTGTCTTTCTCTCTAGCCCCCTCAGGAAAGAGACTATATTTGTACT 240

QY 2822 GTACCTAGGGTCTGGAAGGAAACATGGAATCAGGATTCATAGACTGATAGGCC 2881
|||||
Db 241 GTACCTAGGGTCTGGAAGGAAACATGGAATCAGGATTCATAGACTGATAGGCC 300
|||||
QY 2882 TATCCACAAGGGCCATGACTGGGAAAGGATGGGAGCAGAGGAATTTGGATTTAG 2941
|||||
Db 301 TATCCACAAGGGCCATGACTGGGAAAGGATGGGAGCAGAGGAATTTGGATTTAG 360
|||||
QY 2942 GGTGACGTACGCTACCCCTAACTTTTGGTGGCCCTGGGCGATGCTTGGAGGCCAGACT 3001
|||||
Db 361 GGTGACGTACGCTACCCCTAACTTTTGGTGGCCCTGGGCGATGCTTGGAGGCCAGACT 420
|||||
QY 3002 GTTAAGCAGGCTGCTGGGCTGTTACTCGTCAACACCTCTGCACCTGCTGCTTTGAGA 3061
|||||
Db 421 GTTAACAGGCTGCTGGGCTGTTACTCGTCAACACCTNTGCACCTGCTGCTTTGAGA 480
|||||
QY 3062 C 3062
|
Db 481 C 481
|
RESULT 11
ABL6692/c
ID ABL6692 standard; DNA; 463 BP.
XX ABL6692;
AC
XX
DT 15-MAY-2002 (first entry)
XX
DE Thyroid cancer related gene sequence SEQ ID NO:5329.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
SN WO200194629-A2.
PN
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.

PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 5329; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 463 BP; 116 A; 108 C; 110 G; 129 T; 0 other;

Query Match 13.5%; Score 438.8; DB 24; Length 463;
Best Local Similarity 99.3%; Pred. No. 1.7e-96;
Matches 451; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2790 CCCCCTCAGGAAAGAGGACTATATTTGTACTGTACCTAGGGGTTCTGGAAGGAAAC 2849
|||||
Db 453 CCCCCTCAGGAAAGAGGACTATATTTGTACTGTACCTAGGGGTTCTGGAAGGAAAC 394
|||||
QY 2850 ATGGAATCAGGATTCATAGACTGATAGGCCCTATCCACAAGGCCCATGCTGGGAAAG 2909
|||||
Db 393 ATGGAATCAGGATTCATAGACTGATAGGCCCTATCCACAAGGCCCATGCTGGGAAAG 334
|||||
QY 2910 GTATGGAGCAGAGAGAAATTTGGGATTTTAGGGTCAGCTACGCTCAACCTTAACCTTTT 2969
|||||
Db 333 GTATGGAGCAGAGAGAAATTTGGGATTTTAGGGTCAG-TACGCTCAACCTTAACCTTTT 275
|||||
QY 2970 GGTGGCCTGGGCGATGCTTGGAGCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTAC 3029
|||||
Db 274 GGTGGCCTGGGCGATGCTTGGAGCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTAC 215
|||||
QY 3030 TCGTCACCACTCTGCACCTGCTGTCTTGAGACTCCAGCCCGAGGACGCCACCTG 3089
|||||
Db 214 TCGTCACCACTCTGCACCTGCTGTCTTGAGACTCCAGCCCGAGGACGCCACCTG 155
|||||
QY 3090 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTCGTGTACTGTCTCGGGTCC 3149
|||||

Db 154 CTCCTGAGCCTCCACTATCTCCCTGTGAGCGGTGAATTCGTACTGTCTCGGTCC 95
Qy 3150 ATATATGAATGTGTGAGCAGGTTCTATCTATTTTAAACACAGATGTTTACAAATAAAGAT 3209
Db 94 ATATATGAATGTGTGAGCAGGTTCTATCTATTTTAAACACAGATGTTTACAAATAAAGAT 35
Qy 3210 TATTTCAACACCACCAAAAAAATAAAAAA 3243
Db 34 TATTTCAACACCACCAAAAAAATAAAAAA 1
RESULT 12
AAT25457
ID AAT25457 standard; cDNA to mRNA; 320 BP.
XX AC AAT25457;
XX DT 17-OCT-1996 (first entry)
DE Human gene signature HUMGS07624.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX OS Homo sapiens.
XX PN W09514772-A1.
XX PD 01-JUN-1995.
XX PF 11-NOV-1994; 94WO-JP01916.
XX PR 12-NOV-1993; 93JP-0355504.
XX PA (MATS/) MATSUBARA K.
XX PA (OKUB/) OKUBO K.
XX PI Matsubara K, Okubo K;
XX DR WPI; 1995-206931/27.
XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX PS Claim 1; Page 1835; 2245pp; Japanese.
XX CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-726837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX SQ Sequence 320 BP; 74 A; 75 C; 86 G; 81 T; 4 other;
Query Match 9.6%; Score 313.4; DB 16; Length 320;
Best Local Similarity 98.7%; Pred. No. 4.1e-66;
Matches 314; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 2752 GATCAGCAGGCGTCATTGCTTTCTCTAGCCCTCAGGAAAGAGGACTA 2811

Db 1 GATCAGCAGGCGTCATTGCTTTCTCTAGCCCTCAGGAAAGAGGACTA 60
Qy 2812 TATTTGTACTGTACCTAGGGGTTCTGGAAGGAAACATGAATCAGGATTCATAGAC 2871
Db 61 TATTTGTACTGTACCTAGGGGTTCTGGAAGGAAACATGAATCAGGATTCATAGAC 120
Qy 2872 TGATAGGCCCTATCCACAAGGCCCATGACTGGGAAAAGGTATGGGAGCAGAGGAGAAT 2931
Db 121 TGATAGGCCCTATCCACAAGGCCCATGACTGGGAAAAGGTATGGGAGCAGAGGAGAAT 180
Qy 2932 GGGATTTTAGGTGCAGCTAGCCTCACCTAACTTTTGGTGGCCTGGGGCATGTCTGA 2991
Db 181 GGGATTTTAGGTGCAGCTAGCCTCACCTAACTTTTGGTGGCCTGGGGCATGTCTGA 240
Qy 2992 GGGCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTACTCGTCACCACTCTGCACCTGC 3051
Db 241 GGGCCAGACTGTTAAGCAGGCTCTGCTGGCCTGTTTACTCGTCACCACTCTGCACCTGC 300
Qy 3052 TGTCTTGAGACTCCATCC 3069
Db 301 TGTCTTGAGACTCCATCC 318
RESULT 13
ABL38341
ID ABL38341 standard; cDNA; 313 BP.
XX AC ABL38341;
XX DT 08-APR-2002 (first entry)
XX DE Human colon tumour antigen polynucleotide SEQ ID NO:1930.
XX KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
XX KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX OS Homo sapiens.
XX PN W0200196388-A2.
XX PD 20-DEC-2001.
XX PF 08-JUN-2001; 2001WO-US18557.
XX PR 09-JUN-2000; 2000US-210899P.
XX PR 20-FEB-2001; 2001US-270216P.
XX PA (CORI-) CORIXA CORP.
XX PI Jiang Y, Harlocker SL, Secrist H;
XX WPI; 2002-114514/15.
XX PT Novel isolated colon tumor polynucleotide differentially expressed in
XX colon tumor or colon metastatic tumor and polypeptides encoded by them,
XX useful for inhibiting development of cancer in patient -
XX Claim 1; SEQ ID 1930; 105pp; English.
XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
XX which were isolated from human colon tumour and colon metastatic tumour
XX cDNA libraries. (I) have cytostatic activity and can be used in vaccine
XX production. (I) can be used for stimulating and/or expanding T cells
XX specific for a tumour protein on contact with the T cells. They are also
XX useful for inhibiting the development of cancer in a patient. (I) can be
XX used as probes or primers for nucleic acid hybridisation, for preparing
XX mutant species primers, or primers for use in genetic constructions. (I)
XX can be used in the diagnosis of a colon tumour.
XX SQ Sequence 313 BP; 68 A; 85 C; 85 G; 73 T; 2 other;
Query Match 9.4%; Score 307; DB 24; Length 313;

Best Local Similarity 99.0%; Pred. No. 1.4e-64; Mismatches 1; Mismatches 2; Indels 0; Gaps 0; Matches 308; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2824 ACCCTAGGGGTTCTGGAAGGGAACATGGAATCAGGATTCATAGACTGATAGGCCCTTA 2883
|||||
Db 1 ACCCTAGGGGTTCTGGAAGGGAACATGGAATCAGGATTCATAGACTGATAGGCCCTTA 60

QY 2884 TCACAAAGGCCATGACTGGGAAAGGTATGGAGCAGAGAGAAATGGGATTTAGGG 2943
|||||
Db 61 TCACAAAGGCCATGACTGGGAAAGGTATGGAGCAGAGAGAAATGGGATTTAGGG 120

QY 2944 TGCAGCTAGCTCACCTAACTTTGGTGGCTGGGCGATGCTTGAGGCCAGACTGT 3003
|||||
Db 121 TGCAGCTAGCTCACCTAACTTTGGTGGCTGGGCGATGCTTGAGGCCAGACTGT 180

QY 3004 TAAGCAGGCTCTGCTGGCTGTTTACTCTGCTCACCCTCTGCACCTGCTCTTGAGACT 3063
|||||
Db 181 TAAGCAGGCTCTGCTGGCTGTTTACTCTGCTCACCCTCTGCACCTGCTCTTGAGACT 240

QY 3064 CCATCAGGCCAGGACGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGGTG 3123
|||||
Db 241 CCATCAGGCCAGGACGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGGTG 300

QY 3124 AACTTCGTGTA 3134
|||||
Db 301 AACTTCGTGTD 311

RESULT 14
ABA09160
ID ABA09160 standard; cDNA; 1986 BP.
XX ABA09160;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human rab8 homologue-encoding cDNA, SEQ ID NO:936.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cyostatic; osteoparalic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnary; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-0503800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR P-PSDB; ABB11916.
XX
XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -

XX Claim 1; Page 805; 1963pp; English.
XX
CC Sequences ABB10981-ABBI2330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth factor activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 1986 BP; 519 A; 545 C; 469 G; 453 T; 0 other;

Query Match 7.0%; Score 229; DB 22; Length 1986;
Best Local Similarity 67.0%; Pred. No. 2.8e-45;
Matches 325; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 45 ATGGCGAAGCAGTACGATGTGCTGCGGCTGCTGTCGTCGGGACTCCGGGTGGGC 104
|||||
Db 19 ATGGCGAAGCAGTACGATGTGCTGTCGTCGGGACTCCGGGTGGGC 78

QY 105 AAGACCTGCTGCTGTCGCGCTTTCACCGACACAGAGTTCACCTCCCGACATCTCCACC 164
|||||
Db 79 AAGACCTGCTGCTGTCGCGCTTTCACCGAGCGCTTCAACTCCACTTTTATCTCCACC 138

QY 165 ATCGGTGTTGACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAGTCCGGATACAG 224
|||||
Db 139 ATAGGAATTGACTTTTAAATTTAGGACCATAGAGCTCGATGGCAAGAGATAAATACATGCAG 198

QY 225 ATCTGGGACACTGCAGGCGAGAGAGATACACAGACCATACAAAGCAGTACTATCGGCGG 284
|||||
Db 199 ATATGGGACACACCGGTTCAGAACGGTTTCGGACGATCACACGGCCTACTACAGGGGT 258

QY 285 GCCCAGGGGATATTTTGGTCTATGACATTTAGCAGGAGCGGCTCTTTACAGCAGCATCATG 344
|||||
Db 259 GCAATGGGCATCATGCTGCTGCTACGACATCACCACAGAGAGAGTCTTCGACAAACATCCGG 318

QY 345 AAGTGGGTGACGTGACGTGGATGAGTAGCGCACAGAGCGCTCCAGAGATCCCTTATTGGG 404
|||||
Db 319 AACTGGATTGCGAACATTTGAGGAGCAGCGCTCTGACAGCTCGAAAGATGATACTCGGG 378

QY 405 AATAAGGCTGATGAGGAGCAGAAACGCGAGTGGGAGAGAGAGCAAGGCGCAGCAGTGGCG 464
|||||

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 16:48:24 ; Search time 5516 Seconds
(without alignments)
17184.162 Million cell updates/sec

Title: US-09-817-198A-1
Perfect score: 3257
Sequence: 1 tgcctgctgccgcgcgcag.....aaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2696.6	82.8	190517	9	CNS01DX4	AL139022 Human chr
2	1192.8	36.6	7924	6	AX347040	AX347040 Sequence
3	1192.8	36.6	7924	6	AX348456	AX348456 Sequence
c	1160.8	35.6	7924	6	AX347041	AX347041 Sequence
c	1160.8	35.6	7924	6	AX348457	AX348457 Sequence
5	1000.2	30.7	1054	6	AX399903	AX399903 Sequence
7	866.2	26.6	895	6	AX014147	AX014147 Sequence
8	813	25.0	3139	10	BC027769	BC027769 Mus muscu
9	577.6	17.7	945	10	RATRA15X	M3677 Sprague-Daw
10	477.4	14.7	481	6	AX396088	AX396088 Sequence
c	438.8	13.5	463	6	AX334820	AX334820 Sequence
c	423.6	13.0	162495	2	AC096084	AC096084 Rattus no
13	416.4	12.8	2560	10	BC013790	BC013790 Mus muscu
c	378.4	11.6	257226	2	AC124556	AC124556 Mus muscu
15	307	9.4	313	6	AX341683	AX341683 Sequence
16	234.2	7.2	2048	9	BC002977	BC002977 Homo sapi
17	234.2	7.2	2818	9	AK025165	AK025165 Homo sapi
18	230.6	7.1	1980	9	S53268	S53268 Homo sapien
19	229	7.0	624	9	AF498943	AF498943 Homo sapi
20	229	7.0	660	9	HSWRAB8	X56741 H.sapiens m
21	227.4	7.0	638	10	S53270	S53270 MEL-RAS-rel
22	224.2	6.9	1337	10	BC019990	BC019990 Mus muscu
23	222.6	6.8	760	4	CFRAB8	X56385 Canine rab8
24	218.2	6.7	765	5	DYGORA2	M38391 Discopyge o
25	217.8	6.7	624	6	AX236078	AX236078 Sequence
26	217.8	6.7	1128	9	BC020654	BC020654 Homo sapi
27	217.8	6.7	1161	6	AX236076	AX236076 Sequence
28	217.8	6.7	1265	9	AB038995	AB038995 Homo sapi
29	217.8	6.7	2497	6	AX285074	AX285074 Sequence
30	217.8	6.7	2497	6	AX285080	AX285080 Sequence
31	217.8	6.7	2497	6	AX285089	AX285089 Sequence
32	217.8	6.7	2497	6	AX285091	AX285091 Sequence
33	217.8	6.7	3077	9	AK001111	AK001111 Homo sapi
34	216.2	6.6	3110	9	HSMB04678	AL833365 Homo sapi
35	211.4	6.5	740	10	RNU53475	U53475 Rattus norv
36	203.2	6.2	991	10	RATRA10X	M3677 Sprague-Daw
37	201.8	6.2	897	10	AF035646	AF035646 Mus muscu
c	198.6	6.1	110000	2	AC048347_0	AC048347 Homo sapi
c	198.6	6.1	110000	2	AC048347_1	Continuation (2 of
40	198.6	6.1	302718	2	AC027292	AC027292 Homo sapi
41	198.4	6.1	1029	9	BC000896	BC000896 Homo sapi
42	198.4	6.1	3533	9	AK023223	AK023223 Homo sapi
43	196.8	6.0	3521	9	HSMB01620	AL136650 Homo sapi
44	196.6	6.0	2048	3	AY069671	AY069671 Drosophi
45	196.6	6.0	2210	3	D84347	D84347 Drosophila

ALIGNMENTS

RESULT 1
CNS01DX4
LOCUS Human chromosome 14 DNA sequence BAC R-840119 of library RPCI-11
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL139022
VERSION HTG.
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 190517)
AUTHORS Hellig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,

||||| GTCTTGAGACTCCATCCAGCCCGAGCAGCCACCTGCTCCTGAGCCTCCACTATCTCCC 160258
QY 3113 TGTGAGGGTGAACCTCGTCTACTGTCGTGCGGTGCCATATATGAATTTGTGACGAGGTT 3172
Db 160259 TGTGAGGGTGAACCTCGTCTACTGTCGTGCGGTGCCATATATGAATTTGTGACGAGGTT 160318
QY 3173 CATCTATTTTAAACACAGATGTTTACAAAATAAAGATTATTTCAAACACC 3223
Db 160319 CATCTATTTTAAACACAGATGTTTACAAAATAAAGATTATTTCAAACACC 160369

RESULT 2
LOCUS AX347040 7924 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 2111 from Patent WO0200928.
ACCESSION AX347040
VERSION AX347040.1 GI:18494928
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.

REFERENCE
AUTHORS Olek.A., Piepenbrock.C. and Berlin.K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 2111 03-JAN-2002;
Epigenomics AG (DE)

FEATURES
source location/Qualifiers
1..7924 /organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 2087 a 83 c 2037 g 3717 t
ORIGIN

Query Match 36.6%; Score 1192.8; DB 6; Length 7924;
Best Local Similarity 75.8%; Pred. No. 2.2e-296;
Matches 1476; Conservative 0; Mismatches 472; Indels 0; Gaps 0;

QY 1276 GGTGCTTCTTCCCTTCCCCACCCACCTGTCCTCATCTGCGCATGCGGCTGCTCCCC 1335
Db 1 GGTGCTTCTTCCCTTCCCTTCCCCACCCACCTGTCCTCATCTGCGCATGCGGCTGCTCCCT 60

QY 1336 AGTGACCTGGGAAAGTGGAGCATCGAGGTAGGAGGAAACAGCAACCGGAGCTCCTCGA 1395
Db 61 AGTGATTTGCGAAAGTGGAGTATCGAGGTAGGAGGAAACCGGTAATTAGGCGAGTTTCGA 120

QY 1396 GCCTGGGGCTGCCCTACTCTACCCATTCGCCGACAGAGCTTTGCCCTTGCTGCTGC 1455
Db 121 GTTTGGGGTGTGTTTATTTATTTATTTATTTTCGATTAGAGTTTGTGTTTGTGTTGT 180

QY 1456 CCGCTCGCTCTTTGGGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1515
Db 181 TCGTTGTTGTTTGGGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 240

QY 1516 GAGGGGTGGCAGGATAAAAGTCACTCCATCTCTACCTCCCATGCGAGATGAACACA 1575
Db 241 GAGGGGTGGTAGGATAAAAAGTATTTATTTATTTATTTATTTATTTATTTATTTATTT 300

QY 1576 ATTTCTCCACCTCGCTCCCAATTTAAAGATGTGACCAAGCCCTGCGGTACTCCAG 1635
Db 301 ATTTCTCCACCTCGCTCCCAATTTAAAGATGTGACCAAGCCCTGCGGTACTCCAG 360

QY 1636 GGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCCAGGCAACCATGCATCCCAAGGG 1695
Db 361 GGGTAAGGAGAGTTTGGGGTACTGATATTTGTTAGGTTAATTTATTTTATAAGGG 420

QY 1696 GAGCATTTGGAATGAAGACTAGCTCCTATGATCAGTTAAGCAAGGAGAGAGCTGG 1755
Db 421 GAGTATTTGGAATGAAGACTAGCTCCTATGATCAGTTAAGCAAGGAGAGAGCTGG 480

QY 1756 CCAGGACAGAGTTTGCACAGCAGAGGGGAATCTACCAACAGCAGGCGCTCCTAGGCC 1815
Db 1756 CCAGGACAGAGTTTGCACAGCAGAGGGGAATCTACCAACAGCAGGCGCTCCTAGGCC 1815

Db 481 TTAGGATAGTAGTTTGTATAGTAGAGGGGAATGTAGTAATAGTAGAGGTTTTTTAGGTTT 540
QY 1816 CATCTTCCATTTCTTAGGTAAAGACATTTCCCTCAGACTCCCGAGCGGAGCTAGC 1875
Db 541 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 600
QY 1876 CTAGCCTTTCAGCAACCAAGTTCTCTCTGGGACCAAGTTTATGGGAGAAGCGCAAGAC 1935
Db 601 TTAGTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
QY 1936 TTATCGGGAG 1995
Db 661 TTTATGGGAAG 720
QY 1996 TTTAAGACAAAGCGCTCATCTTTGCCCTCTACCTCTACCTCTAGCTAGGCTTGGAGGCTT 2055
Db 721 TTTAAGATAAAGCGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 780
QY 2056 ACACGTGTGCTACAGGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2115
Db 781 ATATTCGTGTTATAGGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 2116 GTTCTCTTAAACCCATATGSCCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGCAACA 2175
Db 841 GTTTTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 900
QY 2176 AGTAATTTTACCAGTTCTACTGGGGTTCTCTGCCACCGCTCCCAAGGTGGCGAGGCTAGG 2235
Db 901 AGTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 960
QY 2236 AAGAGGTCATCTTAAAGCCACATTTAGCTGCATGCTGCGTGGCTGCAGCCAAACAAAGA 2295
Db 961 AAGAGGTTATTTTAAAGTTATATATTTAGTTGCTGTTTATTTGCTGTTTATTTTAAATTAAGA 1020
QY 2296 ACTGCGTGTGAGTATTCACTAAGAACCAAAATTCAGGCGCTCATATCTGAAGA 2355
Db 1021 ATTGGGTGTGAGTATTATTAATTAAAGATTTAAATTTAGGTATTTATATCTGAAGA 1080
QY 2356 TAAGAACCTCACCTTCTCTCTCCAAAAGAGTGGGAAAGAACCAATCAACCTCTTC 2415
Db 1081 TAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1140
QY 2416 CTCTGACTTACCAACAGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2475
Db 1141 TTTTGTATTTTAAATTTAGGAAATAGTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 2476 GTATAGCTTAGTGTGGAAGCAAG 2535
Db 1201 GTATAGTTTAGTGTGGAAGCAAG 1360
QY 2536 GAAAG 2595
Db 1261 GAAAG 1320
QY 2596 GATGGGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGGCTAGGAGAGAGAGAGAGAG 2655
Db 1321 GATGGGAGTTTGGGTATTAAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1380
QY 2656 GATGCCAGCCCTCAAGTCTCTCTTCAGCCAGGAGCTTGAAGAGTTATTTATTTGGCAGTGGC 2715
Db 1381 GATGTTAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1440
QY 2716 TCCAACTGTGGACAGAGATTTTCCCTGAGAGATCAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2775
Db 1441 TTTAAATTTGGGATTTAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1500
QY 2776 TCTTCT 2835
Db 1501 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1560
QY 2836 CTGGAAGGGAACATGAATCAGAGTTCTATAGACTGATAGGCCCTTATCCCAAGAGGCC 2895
Db 1561 TTGGAAGGGAACATGAATCAGAGTTCTATAGACTGATAGGCCCTTATCCCAAGAGGCC 1620

[illegible]

RESULT 3
 AX348456
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 7924 bp DNA linear
 151 from Patent WO0202806.
 AX348456
 AX348456.1 GI:18614492
 .
 synthetic construct.
 synthetic construct
 artificial sequences.
 1
 Olek.A., Piepenbrock,C. and Berlin,K.
 Method and nucleic acids for pharmacogenomic methylation analysis
 Patent: WO 0202806-A 151 10-JAN-2002;
 Epigenomics AG (DE)
 Location/Qualifiers
 1..7924
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"
 2087 a 83 c 2037 g 3717 t
 BASE COUNT
 ORIGIN

[illegible]

Qy	1576	ATTTCTCTCCACCTGCTGCCCTCCCAAAATTTAAAGATGTGGACCAAGGCCTGTGGGTACTCCAG	1635
Db	301	ATTTTTTTTTTATTTGGTTTTTAAATTTAAAGATGTGGATTAAGGTGTGTGGGTATTTTAG	360
Qy	1636	GGCAGGAGAGCCCTGGGTGAGTGCACACTGTCTAGGCCCAACCATGCACTCCACAAGAGG	1695
Db	361	GGGTAGGAGAGTTTTGGGGTTAGTGCATATCTTTAGGTTAAATATGTATTTTATTAAGAGG	420
Qy	1696	GAGCATTTTGGAAATGAAGCACTAGCTCTTATCTATTCAGGTTAAAGCAAGGAGAGCTGG	1755
Db	421	GAGTATTTGGAAATGAAGGATTAGTTTTTATGTATTAAGTTAAAGGTAAGGAGAGTTGG	480
Qy	1756	CGAGGGACAGAGTTTGCACACAGAGAGGGGAATGTAGCAACAGACAGAGGCCTCTTAGGCC	1815
Db	481	TTAGGATAGTAGTTTGTATGTAGTAGAGGGGAATGTAGTAATAGGTGGTTTTTTTAGTGTT	540
Qy	1816	CATCTTCCATTTCTTAGTTAAGAAGAGCATTTCTTCAGACTCCACAGGGCGAGGACTGAGC	1875
Db	541	TATTTTTTATTTTTTAGTTAAGAAGAGTATTTTTTTTAGATTTTAGCCGAGAGATTGAGT	600
Qy	1876	CTAGCCTTTCAGCAACCAAGGTTCTCTCTGGGACCCAAAGTTTATGGGAGAAGGGCAAGAC	1935
Db	601	TTAGTTTTTAGTAATTAAGSTTTTTTTGGGATTTTAAAGTTTATGGGACAAGGTAAGAT	660
Qy	1936	TTCAATGGGAAGAGAAAGAGCCCTGGGTAGAAACCGTTGGTGCTGTCTCTCTTTGGCC	1995
Db	661	TTTATGGGAAGAGAAAGAGGTTTTTGGGTAGAAACGTTTGGTGTCTGTCTTTTTTTGGTT	720
Qy	1996	TTTAAAGACAAGCGCTCATCTTCCCTCTACCTCTCTGATAGCTGTGAGGGTTGCCAACC	2055
Db	721	TTTAAAGATAAAGCGTTTATTTTGTTTTTTATTTTTTGTATAGGTTTGTAGGGTTGTAAAT	780
Qy	2056	ACACTGTGGCTACAGCTGGAGGGAAGAGCACTCCCTCTCCAGAGTGCATGTTCAGGAA	2115
Db	781	ATATTGGGTTATAGCTGGAGGGAAGAGATTTTTTTTTTAGAGTGTATGTTTAGGAA	840
Qy	2116	GTTFCTTTAAACCCCATATAGCCCAAGAGTAGCTCGTGTAGAGGCCCTTTTAAAGACGGAACA	2175
Db	841	GTTTTTTTTAAATTTTATGTTTTAAGAGTAGTTCGTAGGAGGTTTTTTTTTAAAGACGGAATA	900
Qy	2176	AGTAATTTACAGTTCTACTGGGGTTCTTGCCCAACCGTCCCAAGGTGGCGAGGCCCTAGG	2235
Db	901	AGTAATTTATTTAGTTTTTATTTGGGGTTTTTGTATTATCGTTTTAAGGTGGCGAGGTTTAGG	960
Qy	2236	AAGAGGTCTATCTTAAAGCCACACATAGCTGCATCTGGCTGGCTCGACGCCAAACAAAGA	2295
Db	961	AAGAGGTTATTTTAAAGTTATATATTAGTTGTATTCGTGTGTAGTAAATAAGA	1020
Qy	2296	ACTGGGTGTAGTATTCATCACTAAGAACCAAAATCCAGGGCACTCATATGTGAAGGA	2355
Db	1021	ATTTGGGTGTGAGTATTTTAAATTAAGAAATTTAAATTTAGGGTATTTATATGTGAAGGA	1080
Qy	2356	TAAGAACTCAGTTCTTACTCTCTCAAAAAGAGTGGGGAAGAAGCAACCATCAACCTTTC	2415
Db	1081	TAAGAATTTTATTTTTTATTTTTTTTAAAGAAAGTGGGGAAGAATTTATTAATTTTTTT	1140
Qy	2416	CTCTGACTTTACCAACACAGAAACAGCAGAGAGGGTGGCTCAGGACTTAGGCACAGG	2475
Db	1141	TTTTTGTATTTTAAATTAGGAATAATGTAGAGAGGGTGGTTTAGGATTTAGGGATAGG	1200
Qy	2476	GTATAGCTTAGTGGTGAAGCAAGGAGAGCAGGAAGTTGTAATCACTCGCTTAATGA	2535
Db	1201	GTATAGTTTAGTGGTGAAGGTAAGGAGAGTAGGAAGTTCTAAATTTATTTGGTTAATGA	1260
Qy	2536	GAAAGGAGACAGTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCCTTGAA	2595
Db	1261	GAAAGGAGAGTAGTTAATTTTAGATGAAGTTGTGATTAGGTTGGAGTTGTTTTTTTGAA	1320
Qy	2596	GATGGGACTCTTGGGTATCAAGACCTATGCCACATCACTACACTGGGGCTAGGGAAGTAGGT	2655
Db	1321	GATGGGATTTTTTGGGTTATTAAGATTTATGTATTTATTTGGGGTTAGGAGTAGGT	1380
Qy	2656	GATGCCAGCCCTCAAGTCTCTCTTACGCCAGGGAGCTTGAGAAAGTTATATTGGCAGTGGC	2715

Db	6962	AAAAATCATTTCTAAACACACATTAACCTACACTACGTAACTACACCAAAAAACAAAAAC	6903
Qy	2298	TGGGTGTTGAGTATTCATCAACTAGAACCAAAATCCAGGGCAGCTCATATGTGAAGGATA	2357
Db	6902	TAAATATTAATATTCATCACTCACTAAACCAAAATCCAAACACATCATATATAAAAAATA	6843
Qy	2358	AGAACCTCACTTCTTACTCTCCAAAAAGAGTGGGAAAGAACCATCAAAACCTTTCCT	2417
Db	6842	AAAACCTCACTTCTTACTCTCCAAAAATAAATAAAAAAACCAACATCAAAACCTTTCCT	6783
Qy	2418	CCTGACTTACCAAAACAGAGAAACAGCAGGAGGGTGCTCAGGACTTAGGACAGGGT	2477
Db	6782	CCTAACTTACCAACCAAAACACACAAAAATAAATACTCAAACTTTAAAAACAAAAAT	6723
Qy	2478	ATAGCTTACATGGTGGAAAGCAGGAGAGAGGAGTGTAAATCACGTGGCTAATGAGA	2537
Db	6722	ATAACTTAATAATAAAAAACAAAAACAAAAATAAATACTAACTAATAATAAA	6663
Qy	2538	AAAGGAGACAGCTAAGCTTAGGATGAAGCTGTGACTAGGCTGGAGTTGCTTCTTTGAAGA	2597
Db	6662	AAAAAAACAACCTAAGCTTAAATAAATACTATAAATACTTAAATACTAATAAATACT	6603
Qy	2598	TGGGACTCCTTGGGTATCAAGACCTATGCCACATCACACTGGGCTAGGAACTAGGTGA	2657
Db	6602	TAAACTCCTTAATATCAAAACCTATACCACATCACATAAACTAAAAATAAATAATAA	6543
Qy	2658	TGCCACCCCTCAAGCTGTCTTCAGCAGGGAGCTTGAGAAAGTTATATTGGGCAGTGGCT	2717
Db	6542	TACCAACCCCTCAATCTATCTTCAACCAAAACCTTAAATAATATATAAACAATAAATC	6483
Qy	2718	CAATCTGTGACACAGATTTTCAGCTTTCCTCTGAAGATCAGGAGGGTGCCATTCATTGTC	2777
Db	6482	CAATCTATAACCAATATTTCACTTTCCTTAAATACTAAACAAAAATACCAATTCATTATC	6423
Qy	2778	TTTCTCTCTAGCCCTCAGGAAGAACGAGCTATATTTGTACTGTACCTGAGGGTTCCT	2837
Db	6422	TTTCTCTCTAGCCCTCAAAAAAATAAATACTATATTTATATACCTTAAATAATCT	6363
Qy	2838	GGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGGCCCTATCCACAGGGCCAT	2897
Db	6362	AAAAAAACCAATAAAATCAAAATTTCTATAAATAATAAACCCTATCCACAAAAACCAT	6303
Qy	2898	GACTGGAAAGATGTGGGAGCAGAGAGAAATTTGGGATTTTAGGTGAGCTACGCTCA	2957
Db	6302	AACTAAAAAATAATAAAAAACAAAAAATAAATAATTTAAAAATAACAACCTACGCTCA	6243
Qy	2958	CCCTAAACTTTTGGTGGCCTGGGGCATGTCTTGAGGCCAGAGCTGTTAAGCAGGCTCTGC	3017
Db	6242	CCCTAAACTTTTAATTAACCTTAAACATATCTTAAACCCCAACTATTTAACCAACTCTAC	6183
Qy	3018	TGGCTGTCTTACTCGTCACCACTCTGCACCTGCTCTTTAGAGACTCCATCCAGCCCCAG	3077
Db	6182	TAACCTATTTACTCGTCACCACTCTACACCTACTATCTTTAAACCTCCATCCAAACCCCA	6123
Qy	3078	GCAGGCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTGTGTACTG	3137
Db	6122	ACAGGCCACTTCTCTCAACCTCCACTATCTCCCTATAACGAATAAATTCGTATACTA	6063
Qy	3138	TGTCCTGGGTCCATATATGAATGTGAGCAGGGTTTCATCTATTTTAAACACAGATGTTTA	3197
Db	6062	TATCTGAAATCCATATATAAATTTAACAATAAATTCATCTATTTTAAACACAATAATTTA	6003
Qy	3198	CAAAATAAAGATTTATTTCAACACCACCAAAAAAATAAATAAATAAATAAATAAATAA	3257
Db	6002	CAAAATAAATAATTTTCAAAACCAACCGATATACTACCTAAATAAATCTTTAAAAATAA	5943

RESULT 6
AX399903
LOCUS
DEFINITION Sequence 74 from Patent WO0218424.
ACCESSION AX399903

AX399903 1054 bp DNA linear PAT 06-JUN-2002

VERSION KEYWORDS SOURCE	AX399903.1 human.	GI:21336157
ORGANISM	Homo sapiens	
REFERENCE	Tang, Y.T., Asundi, V., Zhou, P., Xue, A.J., Ren, F., Zhang, J., Wang, J.R., Zhao, Q.A., Wang, D., Liu, C., Drmanac, R.T. and Wehrman, T. Nucleic acids and polypeptides Patent: WO 0218424-A 74 07-MAR-2002; HYSEQ, INC. (US)	
FEATURES source	Location/Qualifiers 1..1054 /organism="Homo sapiens" /db_xref="taxon:9606" 81..719 /note="unnamed protein product" /codon_start=1 /protein_id="CAD33347.1" /db_xref="gi:21336158" /translation="MAKQYDLVRLILLIGDSGVGKTCILCRFTDNEFHSHSTIGVD FKMTIEVDGIKVRIOIWDTAGOERYQTITKOYVRAOGIFLVYDISERSVQHIIMKW VSDVDEYAPGVQKILIGNKADEKROVREGOOLAKKEYGMDRYETSACTNLNIKE SFTRLTELVLQAHKLEGLRMRASNELALAELEEBEGKPEGPANSKTWC"	240 a 314 c 304 g 195 t 1 others
BASE COUNT	240 a 314 c 304 g 195 t 1 others	
ORIGIN		
Query Match	30.7%;	Score 1000.2; DB 6; Length 1054;
Best Local Similarity	99.6%;	Pred. No. 8.2e-247;
Matches 1002;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
QY	2	GCCCGCTGCCCGCCGACGTTCCCGGCCCGCTGGCCCCCAGTCATGCGGAAGCAGTACGA 61
Db	38	GCCCGCTGCCCGCCGACGTTCCCGGCCCGCTGGCCCCCAGTCATGCGGAAGCAGTACGA 97
QY	62	TGTGCTGTTCCGGCTGCTGTCATCGGGACTCCGGGTGGCAAGACCTGCTGCTGTG 121
Db	98	TGTGCTGTTCCGGCTGCTGTCATCGGGACTCCGGGTGGCAAGACCTGCTGCTGTG 157
QY	122	CGCTTTCACCGCAACGAGTTCACCTCCCGACATCTCCACCATCGGTGTTGACTTTAA 181
Db	158	CGCTTTCACCGCAACGAGTTCACCTCCCGACATCTCCACCATCGGTGTTGACTTTAA 217
QY	182	GATGAAGACCATAGAGGTAGCGCATCAAAGTGGCGATACAGATCTGGGACACTGCGAG 241
Db	218	GATGAAGACCATAGAGGTAGCGCATCAAAGTGGCGATACAGATCTGGGACACTGCGAG 277
QY	242	GCAGGAGATACCCAGCCATCACAAAGCAGTACTATCGCGGGCCCGAGGGATATTTT 301
Db	278	GCAGGAGATACCCAGCCATCACAAAGCAGTACTATCGCGGGCCCGAGGGATATTTT 337
QY	302	GCTCTATGACATTAGCAGCGAGCGCTTTACACGACATCATGAAGTGGGTGAGTGACGT 361
Db	338	GCTCTATGACATTAGCAGCGAGCGCTTTACACGACATCATGAAGTGGGTGAGTGACGT 397
QY	362	GATGAGTAGCGCACAGAGCGGTCCAGAGATCTTTATTTGGGAATAAGCTGATGAGGA 421
Db	398	GATGAGTAGCGCACAGAGCGGTCCAGAGATCTTTATTTGGGAATAAGCTGATGAGGA 457
QY	422	GCAGAACCGCAGGTGGGAAGAGAGCAAGGCGACGAGCTGGCCGAAGAGTATGGCATGGA 481
Db	458	GCAGAACCGCAGGTGGGAAGAGAGCAAGGCGACGAGCTGGCCGAAGAGTATGGCATGGA 517
QY	482	CTTCTATGAACAAGTGCCTGCAACCAACCTCAACATTTAAAGAGTCAATTCACGCGTCTGAC 541
Db	518	CTTCTATGAACAAGTGCCTGCAACCAACCTCAACATTTAAAGAGTCAATTCACGCGTCTGAC 577
QY	542	AGAGCTGGTGTGCGAGGCCCATAGGAAGAGCTGGGAAGGCTCCCGATGCGTGCCAGCAA 601
Db	578	AGAGCTGGTGTGCGAGGCCCATAGGAAGAGCTGGGAAGGCTCCCGATGCGTGCCAGCAA 637
QY	602	TGAGTTGGCACTGGCAGAGCTGGGAGGAGGAGGAGGCAAAACCCGAGGCCCGCAACTC 661

```
|||||
Db 638 TGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGCAAAACCCGAGGGCCCGAGAACTC 697
QY 662 TTCGAAAACCTGCCTGCTGAGTCTCTGTGGGGCACCACACAGACACCCCTCTTCCC 721
Db 698 TTCGAAAACCTGCCTGCTGAGTCTCTGTGGGGCACCACACAGACACCCCTCTTCCC 757
QY 722 TCAGGAGCCCGTGGGCAGACAGGGGAGCCGGGGCTTTGCCCTGTCTGCTGCTCTCGTG 781
Db 758 TCAGGAGCCCGTGGGCAGACAGGGGAGCCGGGGCTTTGCCCTGTCTGCTCTCATG 817
QY 782 TGATGACCCCTATTGAGTATCAGTAGCCACTACTCCCTGCTGCTGCTGCTGAGAGGGGCTC 841
Db 818 TGATGACCCCTATTGAGTATCAGTAGCCACTACTCCCTGCTGCTGCTGCTGAGAGGGGCTC 877
QY 842 TGCTGTCATCTCAAGCAGCCCTGTGCCAGCCGCTGTCACCCCTGGAGTGCTCTTCTTCAG 901
Db 878 TGCTGTCATCTCAAGCAGCCCTGTGCCAGCCGCTGTCACCCCTGGAGTGCTCTTCTTCAG 937
QY 902 CCTGTTTCCCAGCCACAGGCTGTCTAGACCCGCCACGATGTGCCGCAAGCACTGTCTCA 961
Db 938 CCTGTTTCCCAGCCACAGGCTGTCTAGACCCGCCACGATGTGCCGCAAGCACTGTCTCA 997
QY 962 CCATCCCGACCCAGCAGACACAGCCAGGGCTGGAGTCCAGGCCA 1007
Db 998 CCATCCCGACCCAGCAGACACAGCCAGGGCTGGAGTCCAGGCCA 1043

RESULT 7
AX014147 AX014147 895 bp DNA linear PAT 07-SEP-2000
LOCUS Sequence 12 from Patent WO9954447.
DEFINITION AX014147
ACCESSION AX014147
VERSION AX014147.1 GI:10040594
KEYWORDS
SOURCE
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 895)
Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarczyk,C.
Human nucleic acid sequences of bladder tumour tissue
PATENT: WO 9954447-A 12 OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source
location/Qualifiers
1..895
/db_xref="taxon:9606"
BASE COUNT 249 a 205 c 228 g 213 t
ORIGIN

Query Match 26.6%; Score 866.2; DB 6; Length 895;
Best Local Similarity 99.5%; Pred. No. 3.1e-212;
Matches 879; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2358 AGAACCTCACTCTCTCTCTCCAAAAGAGTGGGGAAGAACCATCAACCTTCTCT 2417
Db 14 AGAACCTCACTCTCTCTCTCTCCAAAAGAGTGGGGAAGAACCATCAACCTTCTCT 73
QY 2418 CCTGACTTACCAAAACAGGAGAGGAGGCTGGCTCAGGACTTAGGACAGGGT 2477
Db 74 CCTGACTTACCAAAACAGGAGAGGAGGCTGGCTCAGGACTTAGGACAGGGT 133
QY 2478 ATAGCTTAGTGGTGAAGCAAGGAGGAGGAGTGTAAATCACTGGCTAATGAGA 2537
Db 134 ATAGCTTAGTGGTGAAGCAAGGAGGAGGAGTGTAAATCACTGGCTAATGAGA 193
QY 2538 AAAGGAGACAGTAACTCTAGATGAAGCTGTGACTAGGCTGGAGTGTCTTCTTGAAGA 2597
Db 194 AAAGGAGACAGTAACTCTAGATGAAGCTGTGACTAGGCTGGAGTGTCTTCTTGAAGA 253
```

```
QY 2598 TGGGACTCCTTGGGTATCAAGACCTATGCCATCACACTGGGGCTAGGAAGTAGGTGA 2657
Db 254 TGGGACTCCTTGGGTATCAAGACCTATGCCATCACACTGGGGCTAGGAAGTAGGTGA 313
QY 2658 TGCCAGCCCTCAAGTCTCTTTCAGCCAGGGACTTGAGAAGTTATATTGGGCAGTGGCTC 2717
Db 314 TGCCAGCCCTCAAGTCTCTTTCAGCCAGGGACTTGAGAAGTTATATTGGGCAGTGGCTC 373
QY 2718 CAATCTGTGGACCAAGTATTTTTCAGCTTTCCCTGAAAGATCAGCAGGGTGCCTATTCATGTC 2777
Db 374 CAATCTGTGGACCAAGTATTTTTCAGCTTTCCCTGAAAGATCAGCAGGGTGCCTATTCATGTC 433
QY 2778 TTTCTCTCTAGCCCTCAGGAAAGAGGACTATATTGTACTGTACCTTAGGGGTCTT 2837
Db 434 TTTCTCTCTAGCCCTCAGGAAAGAGGACTATATTGTACTGTACCTTAGGGGTCTT 493
QY 2838 GGAAGGAAAACATGGAATCAGGATCTATAGACTGTAGAGCCCTATCCAAAGGGCCAT 2897
Db 494 GGAAGGAAAACATGGAATCAGGATCTATAGACTGTAGAGCCCTATCCAAAGGGCCAT 553
QY 2898 GACTGGGAAAAGGTATGGAGCAGAGGAGATTTGGGATTTTAGGGTGCAGCTACGCTCA 2957
Db 554 GACTGGGAAAAGGTATGGAGCAGAGGAGATTTGGGATTTTAGGGTGCAG-TACGCTCA 612
QY 2958 CCCTAAACTTTTGGTGGCTGGGGCATGCTTTGAGGCCCAGACTGTTAAGCAGGCTCTGC 3017
Db 613 CCCTAAACTTTTGGTGGCTGGGGCATGCTTTGAGGCCCAGACTGTTAAGCAGGCTCTGC 672
QY 3018 TGGCCTGTTTACTCGTCCACCCTCTGCACCTGCTGTCTTGAGACTCCATCCAGCCCCAG 3077
Db 673 TGGCCTGTTTACTCGTCCACCCTCTGCACCTGCTGTCTTGAGACTCCATCCAGCCCCAG 732
QY 3078 GCAGCCACCTGCTCTCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTCTGTA 3137
Db 733 GCAGCCACCTGCTCTCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTCTGTA 792
QY 3138 TGCTCTGGGTCCATATATGAATTGTGACAGGGTTCATCTATTTTAAACACAGATGTTTA 3197
Db 793 TGCTCTGGGTCCATATATGAATTGTGACAGGGTTCATCTATTTTAAACACAGATGTTTA 852
QY 3198 CAAATAAAGATTATTTCAAACCCACCAAAAAA 3240
Db 853 CAAATAAAGATTATTTCAAACCAAAAAA 895

RESULT 8
BC027769 3139 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, clone MGC:38375 IMAGE:5345297, mRNA, complete cds.
DEFINITION BC027769
ACCESSION BC027769
VERSION BC027769.1 GI:20380721
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3139)
Direct Submission
AUTHORS Strausberg,R.
TITLE
JOURNAL
REFERENCE
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
```


Db 1625 CGATGGGTGGCAGGAT--TAAGTACACCTGTGTTCTTCTACCTCCATGCCTC----- 1675
QY 1574 CAATTTCTCTCCACCTGCTCCCAATTTAAAGATGTGGACCAAGCCCTGTGGTACTCC 1633
Db 1676 ----TCTCTCCTACCTGGATTCAAAATTTAAAGATGGGATCGAGTCTGAGGACAC-CA 1730
QY 1634 AGGGCAAGAGAGCCCTGGGTCAGTACACTGTGTCAGGCCCAACCATGCACTCCACAAG 1693
Db 1731 AGGGCAAGGGAGCCCTGGGTCA--GATACTGTATAGTCAAGTGTGAATCCAGATAG 1788
QY 1694 GGGAGATTGGAATTAAGGACTAGCTCTTATGTATAGTTAAGCAAGGGAGAGCT 1753
Db 1789 AGGCAAGCGGATGTCAGGGGGGCTCTTGAACACAGGGTTAGAGCAAGAGCAGAGGG 1848
QY 1754 GGCAGGGACAGCAGTTTCACAGCAGAGGGGAATGTAGCAACAGCAGCGGCTCTTAGGC 1813
Db 1849 AGC-----TTCGCAGACAGAGTGTCTGTGGACTACAGGGCATCAGAGT 1896
QY 1814 CCATCTTCCATTCTTAGGTAAAGAGCATTTCTCAGACTCCAGCGGAGGACTGA 1873
Db 1897 GACAAGTAATC-----AACCACCTCTCTTTCACTGA 1927
QY 1874 GCTACGCTTCAGCAACCAAGTTCTCTTGGNACCCAAAGTTTATGGGAGAGGGCAAG 1933
Db 1928 ACCAGCCTTAGGGAGTTAGGTTTCCCGAGAGTCTGGAACAGGCCTTAAGACAAGGG 1987
QY 1934 ACTTATGGGAGAGAGAGAGGAGGCGCTGGGTAGAAACGCTTGTGCTGTCTTCTTGG 1993
Db 1988 AGAGAGGGTGGTA-----TGATGGAATACTTGTGCTGTCTTCTTGG 2034
QY 1994 CTTTAAAGACAAGCGCTCATCTTGCCCTCTACCTCTGATAGGCTTGAGGGTTGGCCAA 2053
Db 2035 CAG-----GACTCCCTCCCGCTTCCCTTATTTCTAAGTCGGCTTTGCCAA 2082
QY 2054 CCACACTGTGGCTACAGGTGGAGGAGAGAGTCTTCTCAGAGTGCCTATGTTTCAGG 2113
Db 2083 CTGCTGTGGCTTACCAAGGGAGGAAAGAGACACCCCTCTCCAGAGTTCTATGCTCAGG 2142
QY 2114 AAGTTTCTTTAACCCTATATGGCCCAAGAGTAGCTGTAGGAGGCGCTTTAAAGACGGAA 2173
Db 2143 AAGTTTCTTTAACC--TGTGGTTCAAGAGTAGCTTGAAGAGGCGCTTGAACCAAC 2200
QY 2174 CAAATTAATTA-----CCAGTTCTACTGGGTTCTGTCGCCACCGTCCCAAGTGGCGAG 2228
Db 2201 AAACAACAACAACAAGCCACCTGTTAGGTTTCATCTTACATTTCCCAAGTGGGAGGA 2260
QY 2229 GCTAGAGAGAGGTATTTTAAAGCCACACATTAGCTGCATGCTGCTGCAGGCCAA 2288
Db 2261 GCTAGGTAAAGAGAGTCCCAAGATGCATCTCAAGCCTGGGCTGGAGCCAAAGCAGA 2320
QY 2289 ACAAGAAGCTGGGTGT---TGAGTATTATCACTAAGAAC--AAAATCCAGGCACT 2342
Db 2321 GGAAGAACTAGGCACTGACCGAGCATCAAGACCAAGTCTTTCAAATCCAGGCA 2380
QY 2343 CATATGTGAAGATAAGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
Db 2381 GGTGTGATGACAGAGAGCCCAACCTCTTTAGTCTCTTAAAAAGGAGATGGGGTGA 2440
QY 2401 ACCATCAACCTTCTCTCTGCTTACCAACACGAGAAACAGCAGGAGGGTGGCTCA 2460
Db 2441 ATCAATGGGATTTCTCTCTGCTTAGCAAAAGGAAGCAGGCGCAGAGAAAGCAAGCTCA 2500
QY 2461 GGACTTA 2467
Db 2501 GGATTTA 2507

RESULT 9
LOCUS RATRAB15X 945 bp mRNA linear ROD 27-APR-1993
DEFINITION Sprague-Dawley (clone LRB9) RAB15 mRNA, complete cds.
ACCESSION M83679
VERSION M83679.1 GI:206536

KEYWORDS LMW GTP-binding protein.
SOURCE Rattus norvegicus (strain Sprague-Dawley) (library: LAMBDA ZAPII)
adult brain cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 945)
AUTHORS Elferink, L.A., Anzai, K. and Schellier, R.H.
TITLE rab15, a novel low molecular weight GTP-binding protein
specifically expressed in rat brain
J. Biol. Chem. 267 (9), 5768-5775 (1992)
MEDLINE 92210533
PUBMED 1313420
FEATURES
Location/Qualifiers
1..945
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="brain"
/dev_stage="adult"
/tissue_lib="LAMBDA ZAPII"
1..945
/gene="RAB15"
220..858
/gene="RAB15"
/codon_start=1
/product="RAB15"
/protein_id="AAA41995.1"
/db_xref="GI:206537"
translation="NAKQYDVLFRLLIGDSGVKGTCLICRFDNEFHSSHISTIGVD
FKMKTIEVDGIKVRIOIWDTAGERYQITTKOYVRRAGCIFLYVYDISERSYOHIMKW
VSDVDYAPEGVKILIGNKADDEKROVREGOOLAKEYGMDFEYFSACINLIKE
SFRTLTELVLQHRKELDLGRLTCASNELALALEEDEKTEGPANSSKTCWC"
BASE COUNT 222 a 259 c 292 g 172 t
ORIGIN

Query Match 17.7%; Score 577.6; DB 10; Length 945;
Best Local Similarity 88.7%; Pred. No. 1e-137;
Matches 637; Conservative 0; Mismatches 79; Indels 2; Gaps 1;
QY 21 TTCCGCGCCCGCTGGCCCGAGTCATGGCGAAGCAGTACGATGTGCTGTTCGGCTGCTG 80
Db 196 TTCCGCGCCCGCTTGTAGTCCCTGCCATGGCGAAGCAGTACGATGTGCTGTTCGGCTACTG 255
QY 81 CTGATCGGGAGTCCGGGTGGCAAGACCTGCCTGCTGTGCGCTTCACCGACACGAG 140
Db 256 CTGATCGGGAGTCCGGGTGGCAAGACCTGCCTGCTATGCGGCTTCACCGACACGAG 315
QY 141 TTCCACTCTCGCACATCTCCACCATCGGTGTGACTTTAAGATGAAGACCATAGAGGTA 200
Db 316 TTCCACTCTCGCATATCTCCACCATCGGTGTGACTTTAAGATGAAGACCATCGAAGTA 375
QY 201 GACGGCATCAAGTGGGATACAGATCTGGGACACTGCGAGGCGAGAGATACCGAGAC 260
Db 376 GACGGCATCAAGTGGGATACAGATTTGGGACACAGCAGCGGCGAGAGTACCGAGACT 435
QY 261 ATCACAAGCAGTACTATCGCGGGCCCGAGGGATATTTTGGTCTATGACATTACGAGC 320
Db 436 ATCACAAGCAGTACTATCGCGGGCCCGAGGGATATTTTGGTCTATGACATTACGAGC 495
QY 321 GAGCGCTCTTACGACACATCATGAAGTGGGTGAGTGGTGGTGGTGGTGGTGGTGGTGGT 380
Db 496 GAGCGCTCTTACGACACATCATGAAGTGGGTGAGTGGTGGTGGTGGTGGTGGTGGTGGT 555
QY 381 GCGTCCAGAAGATCCTTATTGGGAATAAGGCTGATGAGGAGCAGAAACCGCAGGTGGGA 440
Db 556 GCGTCCAGAAGATCCTTATTGGGAATAAGGCTGATGAGGAGCAGAAACCGCAGGTGGGG 615
QY 441 AGAGACCAAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAACAAGTGGC 500
Db 616 AGAGACCAAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAACAAGTGGC 675

QY 501 TGCACCAACCTCAACATTAAAGAGTCATTCACGGCTGTGACAGAGCTGGTGTGCAGGCC 560
Db 676 TGCACCAACCTTAACATTAAAGAGTCCTTCACTCGTGTGACCGAGCTGGTGTGCAGGCT 735
QY 561 CATAGGAAGAGCTGGAAGGCCCTCCGATCGCTGCCAGCAATGAGTTGGCACTGGCAGAG 620
Db 736 CACAGGAAGAGCTGGATGTCTCCGACATATGTGCCAGCAATGAGCTCGCACCTGGCCGAG 795
QY 621 CTGAGGAGGAGGAGGAGCAACCCAGGCGCCAGCAACTCTTCGAAAACTGCTGTGTGC 680
Db 796 CTGAGGAGGAGGAGGAGCAACCTTAAGGCCAGCAACTCTTCGAAAACTGCTGTGTGC 855
QY 681 TGA--GTCCTGTGTGGGGCACCCACACAGACACCCCTCTTCCTCTCAGAGGCCCGTGG 736
Db 856 TGAGGGTCTGTGTGGAATCTCCCTCCCGCAGTCCCGCTTCCTCTCAGGAGCCCGGTGG 913

RESULT 10
AX396088 . AX396088 481 bp DNA linear PAT 18-MAY-2002
LOCUS Sequence 303 from Patent W00212328.
DEFINITION AX396088
ACCESSION AX396088
VERSION AX396088.1 GI:21066835
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS King, G.E., Meagher, M.J., Xu, J., and Secriest, H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0212328-A 303 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source 1. .481
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 111 a 113 c 130 g 125 t 2 others
ORIGIN

Query Match 14.7%; Score 477.4; DB 6; Length 481;
Best Local Similarity 99.4%; Pred. No. 6.7e-112;
Matches 478; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2582 GTTGCTTCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2641
Db 1 GTTGCTTCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCACACTGGG 60
QY 2642 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGAGCTTGAGAAGTTA 2701
Db 61 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTGTCTTCAGCCAGGAGCTTGAGAAGTTA 120
QY 2702 TATTGGCAGTGGCTCCAATCTGTGGACCAAGTATTCAGCTTTCCCTGGAAGATCAGGCAG 2761
Db 121 TATTGGCAGTGGCTCCAATCTGTGGACCAAGTATTCAGCTTTCCCTGGAAGATCAGGCAG 180
QY 2762 GGTGCCATTCTGCTCTCTCTCTAGCCCTCTCAGAAAGAGGACTATATTGTACT 2821
Db 181 GGTGCCATTCTGCTCTCTCTCTAGCCCTCTCAGAAAGAGGACTATATTGTACT 240
QY 2822 GTACCTTAGGGTTCTGGAAGGAAACATGGAAATCAGGATTCATAGACTGATAGGCC 2881
Db 241 GTACCTTAGGGTTCTGGAAGGAAACATGGAAATCAGGATTCATAGACTGATAGGCC 300
QY 2882 TATCCAAAGGGCATGACCTGGGAAAGGATGGGACGACAGGAAATGGGATTTAG 2941
Db 301 TATCCAAAGGGCATGACCTGGGAAAGGATGGGACGACAGGAAATGGGATTTAG 360
QY 2942 GGTGCAGCTACGCTACCCCTAAACTTTTGTGGCCTGGGCGATGCTTTGAGGCCACACT 3001
Db 361 GGTGCAGCTACGCTACCCCTAAACTTTTGTGGCCTGGGCGATGCTTTGAGGCCACACT 420

QY 3002 GTTAAAGCAGGCTCTGCTGGCGCTGTTTACTCGTCACACACCTCTGCACCTGCTGTCTTGAGA 3061
Db 421 GTTAAACAGGCTCTGCTGGCGCTGTTTACTCGTCACACACCTNTGCACCTGCTGTCTTGAGA 480
QY 3062 C 3062
Db 481 C 481

RESULT 11
AX334820/c AX334820 463 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 5329 from Patent W00194629.
DEFINITION AX334820
ACCESSION AX334820
VERSION AX334820.1 GI:18125539
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D.R., and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature gene sets
JOURNAL Patent: WO 0194629-A 5329 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source 1. .463
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 116 a 108 c 110 g 129 t
ORIGIN

Query Match 13.5%; Score 438.8; DB 6; Length 463;
Best Local Similarity 99.3%; Pred. No. 6.2e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2790 CCCCTCAGGAAAGAGGACTATATTTGTACTGTACCTAGGGGTTCTGGAAGGAAAAAC 2849
Db 453 CCCCTCAGGAAAGAGGACTATATTTGTACTGTACCTAGGGGTTCTGGAAGGAAAAAC 394
QY 2850 ATGGAATCAGGATTTCTAGACTGATAGGCCCTATCCACAGGCCCATGACTGGGAAAG 2909
Db 393 ATGGAATCAGGATTTCTATAGACTGATAGGCCCTATCCACAGGCCCATGACTGGGAAAG 334
QY 2910 GTATGGCAGCAGAGGAGAAATGGGATTTAGGGTGACGTACCTCACCTTAACCTTTT 2969
Db 333 GTATGGGAGCAGAGGAGAAATGGGATTTAGGGTGACG-TACGCTCACCTTAACCTTTT 275
QY 2970 GGTGGCCTGGGCGATGTCTTTCAGGCCCGAGACTGTTAAGCAGGCTCTCTGSCCTGTTTAC 3029
Db 274 GGTGGCCTGGGCGATGTCTTTCAGGCCCGAGACTGTTAAGCAGGCTCTCTGSCCTGTTTAC 215
QY 3030 TCGTCACCACTCTGCACCTGTCTGTGAGACTCCATCCAGCCCCAGGCACGCCACTG 3089
Db 214 TCGTCACCACTCTGCACCTGTCTGTGAGACTCCATCCAGCCCCAGGCACGCCACTG 155
QY 3090 CTCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTCGTGTCTGCTCGGTCC 3149
Db 154 CTCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTCGTGTCTGCTCGGTCC 95
QY 3150 ATATATGAATTTGTGAGCAGGTTTCATCTATTTTAAACACAGATGTTTACAAAATAAGAT 3209
Db 94 ATATATGAATTTGTGAGCAGGTTTCATCTATTTTAAACACAGATGTTTACAAAATAAGAT 35
QY 3210 TATTTCAACACCCAAAAAATAAAAAA 3243
Db 34 TATTTCAACACCCAAAAAATAAAAAA 1

RESULT 12
AC096084/c

LOCUS	AC096084	162495 bp	DNA	linear	HTG 24-AUG-2002
DEFINITION	Rattus norvegicus clone CH230-16122, *** SEQUENCING IN PROGRESS				
VERSION	AC096084	***, 70 unordered pieces.			
KEYWORDS	HTG; HTGS_PHASE1.				
SOURCE	Norway rat.				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 162495)				
AUTHORS	Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,Y., Chen,Z., Chu,J., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Y., Divya,K., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackeleleh,O., Okwou,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.				
TITLE	Direct Submission				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 162495)				
AUTHORS	Worley,K.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	3 (bases 1 to 162495)				
AUTHORS	Rat Genome Sequencing Consortium.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One				
COMMENT	Baylor Plaza, Houston, TX 77030, USA On Aug 9, 2002 this sequence version replaced gi:21723229. ----- Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: GEJZ Center clone name: CH230-16122 ----- Summary Statistics Sequencing vector: Plasmid; Chemistry: Dye-terminator Big Dye: 100% of reads Assembly program: Phrap; version 0.990329 Consensus quality: 116551 bases at least Q40 Consensus quality: 123470 bases at least Q30 Consensus quality: 127758 bases at least Q20 ----- * NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html). * NOTE: This is a 'working draft' sequence. It currently * consists of 70 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * * 1 1247: contig of 1247 bp in length * 1248 1347: gap of unknown length * 1348 2634: contig of 1287 bp in length * 2635 2734: gap of unknown length * 2735 3987: contig of 1253 bp in length * 3988 4087: gap of unknown length * 4088 5288: contig of 1201 bp in length * 5289 5388: gap of unknown length * 5389 6586: contig of 1198 bp in length * 6587 6686: gap of unknown length * 6687 8143: contig of 1457 bp in length * 8144 8243: gap of unknown length * 8244 9249: contig of 1006 bp in length * 9250 9349: gap of unknown length * 9350 10754: contig of 1405 bp in length * 10755 10854: gap of unknown length * 10855 12161: contig of 1307 bp in length * 12162 12622: gap of unknown length * 12622 13639: contig of 1378 bp in length * 13640 13739: gap of unknown length * 13740 14946: contig of 1207 bp in length * 14947 15046: gap of unknown length * 15047 16342: contig of 1296 bp in length * 16343 16443: gap of unknown length * 16443 17772: contig of 1330 bp in length * 17773 17872: gap of unknown length * 17873 19007: contig of 1135 bp in length * 19008 19107: gap of unknown length * 19108 20796: contig of 1689 bp in length * 20797 20896: gap of unknown length * 20897 22145: contig of 1249 bp in length * 22146 22245: gap of unknown length * 22246 24330: contig of 2085 bp in length * 24331 24430: gap of unknown length * 24431 25589: contig of 1159 bp in length * 25590 25689: gap of unknown length * 25690 27349: contig of 1660 bp in length * 27350 27449: gap of unknown length * 27450 28577: contig of 1128 bp in length * 28578 28677: gap of unknown length * 28678 30291: contig of 1614 bp in length * 30292 30391: gap of unknown length * 30392 32126: contig of 1735 bp in length * 32127 33226: gap of unknown length * 33227 33343: contig of 1117 bp in length * 33343 33227				

33444 33443: gap of unknown length
33444 34975: contig of 1532 bp in length
34976 35075: gap of unknown length
35076 37579: contig of 2504 bp in length
37580 37679: gap of unknown length
37680 39003: contig of 1324 bp in length
39004 39103: gap of unknown length
39104 40564: contig of 1461 bp in length
40565 40664: gap of unknown length
40665 41979: contig of 1315 bp in length
41980 42079: gap of unknown length
42080 43598: contig of 1519 bp in length
43599 43698: gap of unknown length
43699 45118: contig of 1720 bp in length
45119 45518: gap of unknown length
45519 47359: contig of 1841 bp in length
47360 47459: gap of unknown length
47460 48918: contig of 1459 bp in length
48919 49018: gap of unknown length
49019 50383: contig of 1365 bp in length
50384 50483: gap of unknown length
50484 52346: contig of 1863 bp in length
52347 52446: gap of unknown length
52447 55257: contig of 2811 bp in length
55258 55357: gap of unknown length
55358 57005: contig of 1648 bp in length
57006 57105: gap of unknown length
57106 59670: contig of 2565 bp in length
59671 59770: gap of unknown length
59771 61952: contig of 2182 bp in length
61953 62052: gap of unknown length
62053 64503: contig of 2451 bp in length
64504 64603: gap of unknown length
64604 65752: contig of 1149 bp in length
65753 65852: gap of unknown length
65853 68166: contig of 2314 bp in length
68167 68266: gap of unknown length
68267 70091: contig of 1825 bp in length
70092 70191: gap of unknown length
70192 72887: contig of 2696 bp in length
72888 72987: gap of unknown length
72988 74782: contig of 1795 bp in length
74783 74882: gap of unknown length
74883 77248: contig of 2366 bp in length
77249 77348: gap of unknown length
77349 79279: contig of 1931 bp in length
79280 79379: gap of unknown length
79380 83158: contig of 3779 bp in length
83159 83258: gap of unknown length
83259 85363: contig of 2105 bp in length
85364 88527: contig of 3064 bp in length
88528 88627: gap of unknown length
88628 90451: contig of 1824 bp in length

Query Match 13.0%; Score 423.6; DB 2; Length 162495;
Best Local Similarity 61.8%; Pred. No. 1.5e-97;
Matches 1227; Conservative 0; Mismatches 664; Indels 95; Gaps 31;

QY 523 AGTCATTACCGCTCTGACAGACTGGTCTGCAGGCCCATAGAGAGGAGCTGGAAGGCC 582
DB 125590 AGTCCTTCACTCGCTGACCGAGCTGGTGCCTGCAGGCTCACAGAAAGAGCTGGATGGTC 125531
QY 583 TCCGGATGGCTGCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGCAAAAC 642
DB 125530 TCCGAACATGTCCAGCAATGAGCTCGACTGGCCGAGCTGGAGGAGAGCAAGGCAAAA 125471
QY 643 CCGAGGCCCGGACGAACTCTTCGAAACCTGCTGCTGCTGA--GTCTGTGTGGGGCACC 700
DB 125470 CTGAAGGCCCGACAACTCTTCAAGACCTTGCTGCTGCTGAGGGTCTCTGTGTGTAATCCT 125411
QY 701 CCACAGCACACCCTCTTCCTCAGAGGCGCCGTGGGCGAGACAGGGGAGCGGGGCTTTG 760

DB 125410 CCCTCCCCCAGTCCCTTCCCTCAGGAGGCCCGCTGGAGGCA---GGAGGCTCTGGGCTTTG 125354
QY 761 CCCTGCTGTGCTCTCTCGGTGATGACCCCTATTGAGTATCAGTAGCCACTACTCTCCCCCT 820
DB 125353 TCCATC-ACTGTCTCTTTCATTGAGATCTCTGTTGAAACCACTAGCTGTACT-CCCT 125296
QY 821 GCCTGCCCTTGAGAGCGGCTCTCTGCTGCACT-CAAGACGCCCTCTGCCACCGCTGCC 879
DB 125295 GCCTGGCCCTGTGAGCAGCTTTGCTGTGGTCTCCGAGCAACCTCTGTCCCCAGCCCTGCT 125236
QY 880 ACCCTGGAGTGGCTTCTTCAGCCTGTTTCCCGCAGCCACAGGCGCTGCTAGACGCCCCACG 939
DB 125235 ATCCTGGAGTGGTAACCTTTTCCACCTGTTTCCAGCGCGAGGCTGCTGTGAGCCCC---- 125180
QY 940 ATGTGCGCGAAGCACTGTCTCACCATCCCGCACCCACCAACAGCAACAGCCAGGCTGGAGT 999
DB 125179 AGAATGTGCTTCGACCATCTCACCAACCCAGGCTCTGGGACAAAGGCTCAAGGCTGGNGT 125120
QY 1000 CCAGGCCACTTTCAGTGTCTCTTCTCCGTGCACTGTCTCTCTCTCTCTCTCTCTCTCTCT 1059
DB 125119 CAAAGCCACTTAAAGACTCCTTTAGTGCATCTCTGTGGGCTCTCTGCTTTTCTCTCTCTT 125060
QY 1060 CTTCGCCCACTTCTCTTCTCTGACCCCTCCCTCCGGTGGCTTTCGTATCAAGACTCCT 1119
DB 125059 CTTCCTCTTCTCCAGTCCCTTCTTCCATGTATTTTGGCTCCACGCTCATCCAGTCTCT 125000
QY 1120 CAAACCCCGTCCCGCTGTGCTGTGTGTCAGCTCGCTTCTTCTCTCTCTCTCTCTCTCTCT 1179
DB 124999 C----CTCCCTAGCTGTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 124943
QY 1180 TATCCAAAGGGATGGACCCAGGCTGTGGGGAGGTTCCACCTTGGATCCAGGAAGAACC 1239
DB 124942 TGCTTAAGGAATGGACCCAGGCTC-----TGCCCCAGGAGGATG 124902
QY 1240 CTCACCCCTGCCTGGGTGGGCCAAAGGCTACAGGGTGTCTTCTCTCTCTCTCTCTCTCTCT 12399
DB 124901 GCTCACCCCTG-CTCTGGGCTGGGCTGAAGGCTA-AGGGTGTCTTCTCTCTCTCTCTCT 124846
QY 1300 CCCACTGTCCCTCATGTGCCATGGGCTGTCTCCAGTGCCTGCCAGTGCCTGCGAAAGTGGAGCATC 1359
DB 124845 CCTGTCTGCCCTTTTGTGCCATGGGCTGTCTCCCAATGGGCTG-GGAGGTAGAACAATT 124787
QY 1360 GAGGTAGGAGGAAACAGCAACCGGGGAGTCTCTGAGGCTTGGGGCTGGGCTTACCTCTTACC 1419
DB 124786 GAGGTAGGAGGAAACACTATGAGTG-GTCTCTCAAGCCT-GGACTGGCCTCTCTCTCTCT 124729
QY 1420 CATTCGCCGACAGAGCTTTGCCCTTGTGGCTGCCCGCTGCTCTCTTGGGGAAGTGA 1479
DB 124728 TTGCCCTTCTCTGGAGTTTGGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 124669
QY 1480 GCTCAGAGGCGAGTGTCTTTCAGAGAAGAAACAAAATGAGGGGTGGCAGGATAAAAAGTC 1539
DB 124668 GTCCAGAGGCGAGCACAGCAGAGAGAAGATGACAAGATGGGTGGGCGGGGATTAAGTCA- 124609
QY 1540 ACCTCCATTTCTACCTCCCATGAGCATGAACACAATTTCTCTCCACTTGGCTTCCAAA 1599
DB 124608 --CTCTGTCTCTTACCTCCCATGAGGAGTGACA--GCCTTTTTCACCCGGGTCTGTGG 124553
QY 1600 TTTTAAAGATGTGCACCAAGGCTGTGGTACTCCAGGGGCAAGSAGAGCCCTGGGGTCAG 1659
DB 124552 ATTTTAAAGATGGGATCAAGGCTTGAAGGCTGAGGAGCGC-CAGAGGGCAAGGGGATCCCCAGGGTCA- 124495
QY 1660 TGACACTGTCCAGCCCAACCATGCATCCACAAAGGGGAGCATTTTGGAAATGAAGGACTAG 1719
DB 124494 -GACACTGTGTAGCTGTGAGCTTGTAGTGTAGGCGCAGAGCGGATGCGAGAGCTGG 124436
QY 1720 CTCCTATGTATCAGGTTAAGAGCAAGGG-AGAGCTGTGCCAGGAGCAGACTTTTGCACAGC 1778
DB 124435 CTCCTAGATGGCAGGTTAGGACAAGAGTAGAGGGGCTTTCACAGAACAAGTGTGCTGT 124376
QY 1779 AGAGGGGAATGTAGCAACAGCAGGGGCTCTCTAGGCCCCATCTTCCATTTCTTGTAGTGAAGA 1838
DB 124375 GGTGACAGCAGGCGGCTCCCTGTGAGGACATTTTACCAGACTCCCTGGCAGAGGTGACA 124316


```
Db 551 GAGAGCAGGGCAGCAGCTGGCTAAGAGTACGGCATGGAGCTTCTACGAAACAAGTGCCT 610
QY 502 GC 503
   ||
Db 611 GC 612

RESULT 14
AC124556/c
LOCUS
DEFINITION Mus musculus chromosome UNK clone RP23-246K11, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
ACCESSION AC124556
VERSION 1 (bases 1 to 257226)
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 257226)
McPherson, J.D. and Waterston, R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 257226)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 257226)
McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Jun 23, 2002 this sequence version replaced gi:21426694.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BA0246K11
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer; 0% of reads
Chemistry: Dye-terminator; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 249197 bases at least Q40
Consensus quality: 250905 bases at least Q30
Consensus quality: 252159 bases at least Q20
Insert size: 214000; agarose-fp
Insert size: 255826; sum-of-ctigs
Quality coverage: 9.36 in Q20 bases; agarose-fp
Quality coverage: 7.92 in Q20 bases; sum-of-ctigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 5033: contig of 5033 bp in length
* 5034 5133: gap of unknown length
* 5134 14755: contig of 9622 bp in length
* 14756 14855: gap of unknown length
* 14856 23882: contig of 8527 bp in length
* 23883 23482: gap of unknown length
* 23483 38499: contig of 15017 bp in length
* 38499 38599: gap of unknown length
```

```
* 38600 46661: contig of 8062 bp in length
* 46662 46761: gap of unknown length
* 46762 57176: contig of 10415 bp in length
* 57177 57276: gap of unknown length
* 57277 86426: contig of 29150 bp in length
* 86427 86526: gap of unknown length
* 86527 104351: contig of 17825 bp in length
* 104352 104451: gap of unknown length
* 104452 122682: contig of 18231 bp in length
* 122683 122782: gap of unknown length
* 122783 169668: contig of 46886 bp in length
* 169669 169768: gap of unknown length
* 169769 244701: contig of 74933 bp in length
* 244702 244801: gap of unknown length
* 244802 246147: contig of 1346 bp in length
* 246148 246247: gap of unknown length
* 246248 248024: contig of 1777 bp in length
* 248025 248124: gap of unknown length
* 248125 251794: contig of 3670 bp in length
* 251795 251894: gap of unknown length
* 251895 257226: contig of 5332 bp in length.
FEATURES
Source
1. .257226
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-246K11"
misc_feature 1..5033
/note="assembly_name:Contig10"
misc_feature 5134..14755
/note="assembly_name:Contig11"
misc_feature 14856..23382
/note="assembly_name:Contig12"
misc_feature 23483..38499
/note="assembly_name:Contig13"
misc_feature 38600..46661
/note="assembly_name:Contig14"
misc_feature 46762..57176
/note="assembly_name:Contig15"
misc_feature 57277..86426
/note="assembly_name:Contig16"
misc_feature 86527..104351
/note="assembly_name:Contig17"
misc_feature 104452..122682
/note="assembly_name:Contig18"
misc_feature 122783..169668
/note="assembly_name:Contig19"
misc_feature 169769..244701
/note="assembly_name:Contig20"
misc_feature 244802..246147
/note="assembly_name:Contig6"
misc_feature 246248..248024
/note="assembly_name:Contig7"
misc_feature 248125..251794
/note="assembly_name:Contig8"
misc_feature 251895..257226
/note="assembly_name:Contig9"
BASE COUNT 68762 a 59441 c 58995 g 68614 t 1414 others
ORIGIN
Query Match 11.6%; Score 378.4; DB 2; Length 257226;
Best Local Similarity 61.2%; Pred. No. 7.8e-86;
Matches 1205; Conservative 0; Mismatches 596; Indels 169; Gaps 30;
QY 523 AGTCATTACGGCTCTGACAGACAGCTGGTGTGAGGCCCTAGAGAGGAGCTGGAAGGCC 582
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135389 AGTCCTTCACTCTGCTGACGGAGCTGGTGTGAGGCCCTAGAGAGGAGCTGATGGTC 135330
QY 583 TCGGATGGTGGCCAGCAATGAGTTGGCACTGCGACAGCTGGAGGAGGAGGCGCAAC 642
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135329 TCCGAACAGCTGCCACAGCAGCTGCACTGGCGGAGGAGGAGGAGGAGGAGGAG 135270
QY 643 CCGAGGGCCCGAGCAACTCTTCGAAAACCTGTGCTGTGAGTCTGTGTGGGGCACCCC 702
```

Db	134265	GC	GGCTCTCTAGAACCAAGG	GTTAGACCAAGAGCAGAGG	GAGC-----TTCCGA	134218
QY	1775	CAG	CAGAGGGGAATGTAGCAACAGCAGG	CGCTCCTAGGCCCATCTTCATTTCTTAGGT	1834	
Db	134217	GA	CAGCGAGTGTGTGGCGACTACAGG	GCATCAGAGTGCACAGTAATC-----	134168	
QY	1835	AAG	AAGAGCATTTCCTCAGACTCCCAGG	CGGAGGACTGACGCTTTCAGCAACCAAG	1894	
Db	134167	-----	AACCACTCCTTTCTAC	TGCAACCGCCTTAGGAGTTAGG	134127	
QY	1895	GT	TCTCTGGGACCCAAAGTTTATGGG	AGNAGGCGAAAGACTTCATGGGAAGAGAGAAG	1954	
Db	134126	GT	TTCCCCAGAGTCTGGCAACAGG	CGCTTAAGACAAGGAAGAGAGGTTGGTA-----	134072	
QY	1955	AAG	GCCTTGGTGAAGACGCTTCGTCTC	TTCTTTGGCTTTTAAGACAAAGCGCTCAT	2014	
Db	134071	-----	TGATGGAAATAC	TCTGGTGTCTTGGCTTCC-----AGACTCC	134032	
QY	2015	CT	TGCCCCCTACCTCCTGTATAG	TCCTTTGAGGGTTTGGCAACCACTGTGGCTACAGTGG	2074	
Db	134031	CT	CCCCCTGGTTCCTTATTTCT	TAAGTCGCTTTTGGCAACTGCTGTGTGGCTACCAAGG	133972	
QY	2075	AGG	AAAGAGACACTCTCTCCAGAGT	GTCTATGTCAGGAAGTTTCTTTAAOCCCATATG	2134	
Db	133971	AGG	AAAGAGACACCTCTCCAGAGT	TTCTATCTCAGGAAGTTTCTTTTAAOCC--TGTT	133914	
QY	2135	GCC	CAAGTAGTCTGTAGGAGCG	CCCTTTAAGACGGAACAGTATTTA-----C	2185	
Db	133913	GT	TCAAGTAGTCTTGAAGGAG	CGCCCTGAAACAACAAACAACAACAAACAAAGC	133854	
QY	2186	CAG	TTCTACTTGGGGTTCTGCC	CACCGTCCCAAGTGTGGCGAGCGCTAGGAAGAGGTC	2245	
Db	133853	CA	CCCTTGGTAGGGTTTCATC	CTTACATCCCAAGTGTGGAGNGCTCAGTAAAGAAGA	133794	
QY	2246	TT	CTTAAGCCACACATPAGTG	TCGTCGTGTGCGGCCAGGCCAAACAAGAACTGGGTGT-	2304	
Db	133793	GT	CCAAAGATGCATTC	CAAGCCTTGGGCTGGAGCCAAAGCAGGAAGAATCTAGGC	133734	
QY	2305	--	TGAGTATTTCATCACTA	AGAACCC--AAATCCAGGCACTCATCTGTGAAGNTAG	2359	
Db	133733	ACC	GAGCATTC	CAAGGACCAAGTGCCTTTCAATTC	CCAGGCGACGCTGTGATGACAGAGAG	133674
QY	2360	AAC	CTCACTTCTTACTCTCC	TCCCAAAAA--GAAGTGGGGAAGAACCATCAAACTTCCT	2417	
Db	133673	CCC	CACACTCTTTGTCTCT	TAAAAAGGAGATGGGGTGAATCAATGGGGATTCTTC	133614	
QY	2418	CCT	GACTTACCAACCAAGG	AAAAACAGCAGGAGAGGCTCAGGACTT	2467	
Db	133613	CCT	GACTTAGCAACAAAGG	AAAGCAGCGCAAGAGAAGCTCAGGATTTA	133564	

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2002, 20:40:59 ; Search time 1956 Seconds
(without alignments)
1755.339 Million cell updates/sec

Title: US-09-817-198A-2
Perfect score: 1105
Sequence: 1 MAKQYDLVRLLLIGDSGVG.....LEEEGRKEPGANSSKTCWC 212

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09817198/runat_13112002_135207_7044/app_query.fasta_1.391
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198_ACGN_1_1_899 @runat_13112002_135207_7044 -ICPU=3
-NO_XLPYX -NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_Other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	996	90.1	1091	12	BF160330	BF160330 601768601
2	933.5	86.3	932	12	BF535642	BF535642 602054039
3	922	83.4	916	13	BF1648588	BF1648588 603277781
4	914	82.7	992	12	BF101730	BF101730 601753464
5	910	82.4	902	12	BF178163	BF178163 601809083
6	885	80.1	699	13	BF557933	BF557933 603237549
7	858.5	77.7	796	12	BF966292	BF966292 602286692
8	838	75.8	1100	12	BF181167	BF181167 601808830
9	829	75.0	616	12	BG247902	BG247902 602359819
10	816	73.8	1121	12	BF182001	BF182001 601805936
11	799	72.3	814	12	BG762967	BG762967 602735022
12	792	71.7	660	13	BG974411	BG974411 602844184
13	777	70.3	911	13	BF1649317	BF1649317 603278051
14	756	68.4	618	12	BF150465	BF150465 uy86b05.y
15	733	66.3	1767	13	BM462147	BM462147 AGENCOURT
16	727	65.8	485	12	BE913096	BE913096 601668580
17	706	63.9	667	13	BF1685808	BF1685808 603309467
18	704	63.7	843	12	BG865609	BG865609 602783763
19	693.5	62.8	577	12	BG703353	BG703353 602685115
20	686	62.1	465	12	BF565114	BF565114 UI-R-B01-
21	670	60.6	479	12	BF320679	BF320679 uz55d10.y
22	666	60.3	543	12	BG019552	BG019552 daa21e04.y
23	663	60.0	539	12	BF022618	BF022618 uy51f07.y
24	658.5	59.6	693	10	BE287683	BE287683 601093322
25	651	58.9	823	13	BM017178	BM017178 603643780
26	649	58.7	385	14	BM936454	BM936454 UI-M-BH2.
27	589	53.3	956	14	BQ944882	BQ944882 AGENCOURT
28	568.5	51.4	704	13	BJ057692	BJ057692 BJ057692
29	554.5	50.2	714	13	BJ259835	BJ259835 602971485
30	551.5	49.9	644	14	BQ387194	BQ387194 NISC-mm22
31	547	49.5	611	13	BJ035681	BJ035681 BJ035681
32	547	49.5	788	13	BI395002	BI395002 pppin.pk0
33	547	49.5	788	9	AUI25167	AUI25167 AUI25167
34	545	49.3	613	14	BM791355	BM791355 K-EST0071
35	545	49.3	635	9	AUI34059	AUI34059 AUI34059
36	545	49.3	676	9	AUI33710	AUI33710 AUI33710
37	545	49.3	686	12	BG709122	BG709122 602675387
38	545	49.3	816	9	AUI19700	AUI19700 AUI19700
39	545	49.3	835	12	BG425680	BG425680 602452987
40	545	49.3	860	9	AUI22740	AUI22740 AUI22740
41	545	49.3	877	13	BI763900	BI763900 603049666
42	545	49.3	897	14	BQ671010	BQ671010 AGENCOURT
43	545	49.3	986	13	BM475359	BM475359 AGENCOURT
44	545	49.3	1032	13	BM459427	BM459427 AGENCOURT
45	545	49.3	1059	14	BM923211	BM923211 AGENCOURT

ALIGNMENTS

RESULT 1
BF160330
LOCUS
DEFINITION BF160330 1091 bp mRNA linear EST 30-OCT-2000
601768601F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3987745 5',
mRNA sequence.
ACCESSION BF160330
VERSION BF160330.1 GI:11040541
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1091)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9195 row: b column: 02
 High quality sequence stop: 654.
 Location/Qualifiers
 1..1091
 /organism="Mus musculus"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:3987745"
 /clone_lib="NCI_CGAP_Lu29"
 /tissue_type="spontaneous tumor, metastatic to mammary."
 Stem cell origin."
 /lab_host="DHI0B"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT	297 a	320 c	317 g	157 t
ORIGIN				
Alignment Scores:				
Pred. No.:	6,93e-117	Length:		1091
Score:	996.00	Matches:		198
Percent Similarity:	94.42%	Conservative:		5
Best Local Similarity:	92.09%	Mismatches:		9
Query Match:	90.14%	Indels:		3
DB:	12	Gaps:		0

US-09-817-198A-2 (1-212) x BF160330 (1-1091)

Db	556	AGTCCTTCACTCGTCTGACGGAGCTGGTGTCTGACGGCCACAGGAAGACCTGGATGGTC	615
Qy	180	euArgMetArgAlaSer-AsnGluLeuAlaLeuAlaGluLeuGluGluGlyLys	199
Db	616	TCGGAACACGTGCGAGCAACAGAGCTGCGCATCGGGCGAGTGGAGGACGAAGCAAA	675
Qy	200	ProGluGlyProAlaAsnSerSerLysThrCysTrpCys	212
Db	676	CCCTGGGGCCACGAATCTTCAAGGACGTGCTGGTGC	714
RESULT 2			
LOCUS	BF535642	602054039F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4193456	5', linear EST 11-DEC-2000
DEFINITION		mRNA sequence.	
ACCESSION	BF535642		
VERSION	BF535642.1	GI:11623010	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE	1 (bases 1 to 932)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabps-femail.nih.gov		
	Tissue Procurement: Jeffrey E. Green, M.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Plate: L1AM9525 row: m column: 09		
	High quality sequence stop: 685.		
FEATURES	Location/Qualifiers		
source	1..932		
	/organism="Mus musculus"		
	/strain="FVB/N"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:4193456"		
	/lab_host="DH10B (T1 phage-resistant)"		
	/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."		
BASE COUNT	239 a 250 c 294 g 149 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1..54e-111	Length:	932
Score:	953.50	Matches:	187
Percent Similarity:	94.63%	Conservative:	7
Best Local Similarity:	91.22%	Mismatches:	9
Query Match:	86.29%	Indels:	2
DB:	12	Gaps:	1
US-09-817-198A-2 (1-212) x BF535642 (1-932)			
Qy	1	MetalLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly	20
Db	145	ATGGCGAAACAGTACGATGTCTGTCGGGTACTGTGTGATCGGGGACTCCGGGGTTGGC	204
Qy	21	LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr	40
Db	205	AAGACATGCCCTGCTGCGCGCTTCCACGACACAGAGTTCACCTCTCGCATATCTCCACC	264
Qy	41	IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln	60
Db	265	ATCGGTGTGACTTAAAGATGAACAGCATTCATGCTAGACGCATCAAGTCAGAAATACAG	324

Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 167 a 154 c 179 g 115 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 8.89e-108 Length: 616
Score: 922.00 Matches: 180
Percent Similarity: 98.92% Conservative: 3
Best Local Similarity: 97.30% Mismatches: 2
Query Match: 83.44% Indels: 1
DB: 13 Gaps: 0

US-09-817-198a-2 (1-212) x BI648588 (1-616)

Qy 28 PheThrAspAsnGluPheHisSerHisIleSerThrIleGlyValAspPheLysMet 47
Db 25 TTCACGACAAAGAGTTCACACTCCCTCGCATATCTCCACCATCGGTGTTGACTTTAAGATG 84
Qy 48 LysThrIleGluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGln 67
Db 85 AAGACTATCGATGTAGACGGCATCAAGTGAGAAATACAGATTGTGGACACACAGGGGAG 144
Qy 68 GluArgTyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlyIlePheLeuVal 87
Db 145 GAGAGTACCAGACTATCACAAAGCAGTACTATCGCGAGCCCA-GGAATATTTTAGTC 203
Qy 88 TyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAsp 107
Db 204 TACGACATTAGCAGTGAGCGCTCCTATCAGCATATCATCAAGTGGTCACTGACGTGGAT 263
Qy 108 GluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGln 127
Db 264 GAGTACGCTCCAGAGAGGTCCAGAAAGATCCTTAATTGGGAATAGGCTGATGAAGAGCAG 323
Qy 128 LysArgGlnValGlyArgGluGlnGlyGlnLeuAlaLysGluTyrGlyMetAspPhe 147
Db 324 AAACGGCAGGTGGGAGAGAGCAGGGCAGCAGCTGGCTAAGGAGTACGGCATGACATTC 383
Qy 148 TyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGlu 167
Db 384 TACGAAACAAGTGCCTGCACCAACCTTAATATTAAGAGTCTCTCACTCGTCTGACGGAG 443
Qy 168 LeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGlu 187
Db 444 CTGGTGTCTGACGGCCACAGAAAGAGCTGGATGTCTCCGAAACACGTGCCACACAGAG 503
Qy 188 LeuAlaLeuAlaGluLeuGluGluGlyLysProGluGlyLysProAlaAsnSerSer 207
Db 504 CTCGGACTGGCGAGCTGGAGGAGGAGCAGGAGCAAACTGAGGGCCACCAACTCTTCA 563
Qy 208 LysThrCysTrpCys 212
Db 564 AAGACCTGCTGTGTC 578

RESULT 4
BF101730
LOCUS 601753464F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3981183 5',
DEFINITION mRNA sequence.
ACCESSION BF101730
VERSION BF101730.1 GI:10884256
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 992)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

Qy 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 325 ATTTGGGACACAGCAGGAGAGGTACCAGACTATCAACAAGCAGTACTATCGGCA 384
Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 385 GCCCAGGAATATTTTATGCTACGACATTAGCAGTACGCGCTCCTATCAGCATATCATG 444
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuGly 120
Db 445 AAGTGGGTACGAGCTGGATGAGTACGCTCCAGAAAGAGTCCAGAAAGATCCTAATTGGG 504
Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
Db 505 RATAAGGCTGATGAAGACAGCAACGCGAGGTGGGAGAGCAGCGGCAGCTGGCT 564
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 565 AAGGAGTACGGCATGGACTTCTACGAAACAAAGTGCTCGCACAACCTTAATATTTCAAGAG 624
Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
Db 625 TCCTTCACTCGTCTGACGAGCTGGTGTGACGGCCACAGAAAGAGCTGGATGGGCTC 684
Qy 181 ArgMetArgAla-SerAsnGluLeuAlaLeuAla---GluLeuGluGluGluGly 199
Db 685 CGAACACGTGCGCGCAGCAGCTCGCATGGCGGAGAGCTGTGAGAGCGGAGCAAGCGCA 744
Qy 199 sProGluGlyPro 203
Db 745 ACCTGATGGCCA 757

RESULT 3
BI648588
LOCUS 603277781F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5317917 5',
DEFINITION mRNA sequence.
ACCESSION BI648588
VERSION BI648588.1 GI:15562824
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 616)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL1806 row: a column: 22
High quality sequence start: 2
High quality sequence stop: 614.
Location/Qualifiers

FEATURES
source 1. 616
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone_image="IMAGE:5317917"
/clone_lib="NCI_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9177 row: p column: 16
High quality sequence stop: 706.
Location/Qualifiers
1. .992
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 251 a 272 c 306 g 163 t
ORIGIN
Alignment Scores:
Pred. No.: 2,01e-106 Length: 992
Score: 914.00 Matches: 181
Percent Similarity: 86.8% Conservative: 4
Best Local Similarity: 84.9% Mismatches: 1
Query Match: 82.7% Indels: 27
DB: 12 Gaps: 1

US-09-817-198a-2 (1-212) x BF101730 (1-992)

QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
|||||
DB 150 ATGGCGAAGACAGTACGATGTGCTGTTCCGGCTACTCTGATCGGGGACATCCGGGGTGGC 209
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeSerThr 40
|||||
DB 210 AAGACATGCTGCTGTGCGCTTCACCGACACGAGTCCACTCTCCGATATCTCCACC 269
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspClyLysValArgIleGln 60
|||||
DB 270 ATCGGTGTTGACTTTAAGATGAAGACTATCGATGTAGCGGCATCAAGTGAGATACAG 329
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
|||||
DB 330 ATTTGGACACACAGCGGAGAGAGTACCAGACTATCACAAAGCAGTACTATCGCGCA 399
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
|||||
DB 390 GCCCAG----- 395
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
|||||
DB 396 -----TACGCTCCAGAGAGGAGTCCGAGAGATCCCAATTTGGG 431
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnLeuAla 140
|||||
DB 432 AATAAGGCTGATCAAGACAGAAACGCGAGTGGGAGAGAGCGGCGGAGCAGCTGGCT 491
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
|||||
DB 492 AAGGAGTACGGCATGGACTTCTACGAACAAGTGCCTGCACCAACCTTATATTAAGAG 551
QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
|||||
DB 552 TCCTTCACCTGCTGACGGAGAGCTGGTGCTGCAGGCCCCACAGGAAGAGCTGGATGCTC 611

QY

181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu-GluGluGlyLysPr 200
|||||
DB 612 CGAACACGTGCCAGCAACGCTCGCACTGGCGAGTGGAGGACGAGGAGGCAACC 671
QY 200 oGluGlyProAlaAsnSerSerLysThrCysTrpCys 212
|||||
DB 672 TGAGGGTCCAGCAAACTCTTCAAAGACCTGCTGGTGC 708

RESULT 5

BF178163 902 bp mRNA linear EST 31-OCT-2000
LOCUS 601809083F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4039701 5',
DEFINITION mRNA sequence.

ACCESSION

BF178163
VERSION BF178163.1 GI:11056305
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE

1 (bases 1 to 902)
NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov

FEATURES

source
1. .902
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4039701"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT

216 a 247 c 258 g 181 t
ORIGIN

Alignment Scores:

Pred. No.: 5.63e-106 Length: 902
Score: 910.00 Matches: 177
Percent Similarity: 97.28% Conservative: 2
Best Local Similarity: 96.20% Mismatches: 4
Query Match: 82.35% Indels: 1
DB: 12 Gaps: 0

US-09-817-198a-2 (1-212) x BF178163 (1-902)

QY 30 AspAsnGluPheHisSerHisIleSerThrIleGlyValAspPheLysMetLysThr 49
|||||
DB 14 GACAAGGTTCACACCTCCGATATCTCCACCATCGGTGTGACTTTAAGATGAAGACT 73
QY 50 IleGluValAspGlyLysValArgIleGlnIleTrpAspThrAlaGlyGlnGluArg 69
|||||
DB 74 ATCGAAGTAGACGGCATCAAAGTGATAATACAGATTTGGGACACAGCAGCGGAGGAGG 133

Qy 70 TyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlyIlePheLeuValTyrAsp 89
 Db 134 TACCAGACTATCACAAAGCAGTACTATCGCGAGCCAGGGAATATTTTAGTCTACGAC 193

Qy 90 IleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyr 109
 Db 194 ATTAGCAGTGGCGCTCTATCAGCATATCATGAAGTGGTCAAGTGGATGAGTAC 253

Qy 110 AlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGluGlnLysArg 129
 Db 254 GCTCCAGAGGAGTCCAGAAATCTTAATGGGATAAGGCTGATGAAGACAGAAACGG 313

Qy 130 GlnValGlyArgGluGlnGlnGlnLeuAlalys-GluTyrGlyMetAspPheTyrG1 149
 Db 314 CAGGTGGGAGAGACAGCGGACAGCTGGCTACCGAGTACGGCATGGACTTCTACGA 373

Qy 149 uThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuVa 169
 Db 374 ACAAAGTGCCTGCACCAACCTTAATATAAGAGTCTTCACTCGTCTGACGGAGCTGGT 433

Qy 169 IleuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAl 189
 Db 434 GCTGACGCCCCACAGCAGAGCTGGTGGTCTCGAACACGTGCCACACAGGACTCGC 493

Qy 189 aLeuAlaGluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysTh 209
 Db 494 ACTGGCCGAGCTGGAGGAGGAGGCAAGCAACCTGAGGGCCACGCAAACTCTTCAAAGAC 553

Qy 209 rCysTrpCys 212
 Db 554 CTGCTGGGTGC 563

RESULT 6
 BI557933 603237549f1 NCI_CGAP_Mam3 Mus musculus cdna clone IMAGE:5290284 5',
 LOCUS mRNA sequence.
 DEFINITION BI557933.1 GI:15445247
 VERSION EST.
 KEYWORDS house mouse.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapb-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM11734 row: b column: 13
 High quality sequence stop: 698.
 Location/Qualifiers
 1. 699
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5290284"
 /clone_lib="NCI_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH

Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999).
 BASE COUNT 164 a 188 c 219 g 128 t
 ORIGIN

Alignment Scores:
 Pred. NO.: 6.1e-103 Length: 699
 Score: 885.00 Matches: 172
 Percent Similarity: 98.87% Conservative: 3
 Best Local Similarity: 97.18% Mismatches: 1
 Query Match: 80.09% Indels: 1
 DB: 13 Gaps: 0

US-09-817-198A-2 (1-212) x BI557933 (1-699)

Qy 1 MetAlaLysGlnTyrAspValLeuPheLeuLeuLeuLeuLeuLeuLeuLeuLeuValGly 20
 Db 156 ATGGCGAAACAGTACGATGTCTGTCGGCTACTGCTGATCGGGGACTCCGGGGTTGGC 215

Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLysSerThr 40
 Db 216 AAGACATGCTGCTGTCGCGCTTCACCGACAAGAGTTCCACTCTCGCATATCTCCACC 275

Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
 Db 276 ATCGGTGTTGACTTTAAGATGAAGACTATCGATGTAGACGGCATCAAAAGTCAGAATACAG 335

Qy 61 IleTTPAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
 Db 336 ATTTGGGACACAGCAGGCGAGAGAGTACCAGACTATCACAAAGCAGTACTATCGGGCA 395

Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
 Db 396 GCCCAGGGAATATTTTAGTCTACGACATAGCAGTGGCGCTCTATCAGCATATCATG 455

Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuLeuGly 120
 Db 456 AAGTGGTCACTGAGTGGATGAGTACGCTCCAGAAAGAGTCCAGAAGATCCTAATATGGG 515

Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnVal-GlyArgGluGlnGlnGlnLeuAl 140
 Db 516 AATCAGGTGATGAAGAGCAAAACGGCAGGTGGGCGAGAGAGAGGCGCAGCAGCTGGC 575

Qy 140 aLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysG1 160
 Db 576 TAAGAGTACGCGATGGACTTCTAGAAACAAGTCCCTGCACCAACCTTAATATTAAGA 635

Qy 160 uSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGlu 176
 Db 636 GTCTTTCACCTGCTGACGGAGTGGTGTGTCAGGCGCCACAGGACACAG 684

RESULT 7
 BI966292 796 bp mRNA linear EST 23-JAN-2001
 LOCUS 602286592f1 NIH_MGC_95 Homo sapiens cdna clone IMAGE:4375366 5',
 DEFINITION mRNA sequence.
 ACCESSION BF966292
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapb-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

NOT ON

```

QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
Db 220 AAGACATGCTGCTGTGCGCTTCAACGCAACAGAGTCCACTCTCGCATATCTCCACC 279
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 280 ATCCGCTGTTGACTTTAAGATGAAGACTATCGATAGACGCAATCAAGAGTGAAGATACAG 339
QY 61 IleTrpAspPheAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 340 ATTTGGGACACAGCAGGCGAGAGGTACGACTATCAAAAGCAGTACTATCTCGCGGA 399
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 400 GCCCAGGGAATATTTTGTAGTCTACGACATTAGCAGTGGCGCTCTATCAGCATATCATG 459
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 460 AAGTGGGTGACGTGACGTGATGAGTACGCTCCAGAGAGGTCCAGAAAGATCTTAATGGG 519
QY 121 AsnLysAlaAspGluGlnLysArgGlnVal-GlyArgGluGlnGlyGlnGlnLeuAl 140
Db 520 AATAAGGCTGATGAAGAGCAGCAACGCGAGGTGGGGAGCAGCAGGCGGCGAGCTGGC 579
QY 140 aLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsn--IleLys 159
Db 580 TAAGGATGACGATGACGTCTACGAAACAAAGTGCCTGGGAACACCTTAATATTACAC 639
QY 160 GluSerPheThrArg-LeuThrGluLeuValLeuGlnAlaHisArgLys-GluLeuGluG 179
Db 640 GAGTCTTCACTGCTCTTGAGGAGCTGCTGTCAGGCCCCACAGGGAACGAGCTGGACT 699
QY 179 lyLeuArgMetArgAlaSerAsnGluLeu-----AlaLeuAlaGluL 193
Db 700 GG-----GCTCCGACACAGGTGGGCCACAGCAGCAGCGCCGCTGGCGGAAC 747
QY 193 eu-----GluGluGluGluGlyLysProGluGly-----ProAlaAsnSers 207
Db 748 TTGGGAGGACGACGACAAAGGAACCCGAGGGGGCCCGCACAACCTTCCCAAGCA 807
QY 207 er 207
Db 808 CC 809

RESULT 9
BG247902 616 bp mRNA linear EST 13-FEB-2001
LOCUS 602359819f1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488197 5',
DEFINITION mRNA sequence.
ACCESSION BG247902
VERSION BG247902.1 GI:12757717
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10334 row: f column: 06
High quality sequence stop: 616.
Location/Qualifiers

FEATURES

```

```

source 1. .616
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488197"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 165 a 152 c 183 g 116 t
ORIGIN

Alignment Scores:
Pred. No.: 7.73e-96 Length: 616
Score: 829.00 Matches: 158
Percent Similarity: 99.38% Conservative: 3
Best Local Similarity: 97.53% Mismatches: 1
Query Match: 75.02% Indels: 0
DB: 12 Gaps: 0

US-09-817-198A-2 (1-212) x BG247902 (1-616)
QY 51 GluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGlnGluArgTyr 70
Db 3 GATGTAGCGCATCAAAAGTGAGATACAGATTTGGGACACAGCAGGCGAGGAGGTAC 62
QY 71 GlnThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuValTyrAspIle 90
Db 63 CAGACTATCAAAAGCAGTACTATCGGCGAGCCCGAGGAATATTTTGTCTACGACATT 122
QY 91 SerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyrAla 110
Db 123 AGCAGTGGCGCTCTCTATCAGCATATCATGAAGTGGGTCAGTACGTCGATGAGTACCT 182
QY 111 ProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArgGln 130
Db 183 CCAGAGGAGTCCAGAGATCTTAATGGGAATAAGGCTGATGAAGACAGCAAGCGGAC 242
QY 131 ValGlyArgGluGlnGlyGlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGluThr 150
Db 243 GTGGGAGAGAGCAGGCGCAGCTGGCTAAGGAGTACGCGTACGCTTCTACGAAACA 302
QY 151 SerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuValLeu 170
Db 303 AGTGCCTGCACCAACCTTAATTAAGAGAGTCTCTCTGACGAGCTGGTGTCTG 362
QY 171 GlnAlaHisArgLysGluLeuGluGluArgMetArgAlaSerAsnGluLeuAlaLeu 190
Db 363 CAGCCCAACAGAAAGAGCTGGATGGTCTCGAACACAGCTGCCAGCAGCTCGCATG 422
QY 191 AlaGluLeuGluGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysThrCys 210
Db 423 GCCGAGCTGGAGGAGGAGGAGCAAGCAACCTGAGGCCCAAGCAACTCTTCAAGACCTGC 482
QY 211 TrpCys 212
Db 483 TGGTGC 488

RESULT 10
BF182001 1121 bp mRNA linear EST 31-OCT-2000
LOCUS 601805936f1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4036664 5',
DEFINITION mRNA sequence.
ACCESSION BF182001
VERSION BF182001.1 GI:11060143
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1121)
AUTHORS NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
robin.humphreys@lthar.com
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM9312 row: h column: 09
 High quality sequence stop: 712.

FEATURES
SOURCE

```
location/qualifiers
1. 1121
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4036664"
/clone_lib="NCI_CGAP_Mam
/tissue_type="tumor, gro
/dev_stage="7 months"
/lab_host="PH10R"
```

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NIH"

	NIN		
BASE COUNT	274 a	328 c	340 g
ORIGIN	178 t	178 t	1 others

Alignment Scores:	
Pred. No.:	9, 29e-94
Score:	816.00
Percent Similarity:	85.32%
Best Local Similarity:	82.63%
Query Match:	73.85%
DB:	12
Length:	1121
Matches:	176
Conservative:	7
Mismatches:	25
Indels:	6
Gaps:	2

US-09-817-198A-2 (1-212) x BF182001 (1-1121)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
 196 ATGGCGAACAAGTACGATGTGCTGTTCGGCTACTGCTGATCGGGGACATCCGGGGTTGGC 255

Qy 21 LysThrCysLeuLeuCysArgPheThrAspIasnGluPheHisSerSerHisIleSerThr 40
 |||||
Db 256 AGACATGCCTGCTGTCCGGTTCACCGCACACCGAGTTCCACGCCCGCATATCTCCACC 315

Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
316
317

Qy 61 I l e T r p A s p T h r A l a G l y G l n G l u - A r g T y r C l n T h r I l e T h r L y s G l n T y r T y r A r g A r 80

Qy 80 gAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMe 100

DD	436	AGCCCCAGGGGAATATTTT	TAGTCTACGACATTTAGCAGT	GAGCGCTCTATCAGCATATCAT	499
QY	100	tLysTrpValSerAspValAspGluTyrAlaProGluClyValClnLysIleLeuIleG1			120

Db	496	GAAGTGGGTCAAGTCACGTGGATGAGTACGCTCCAGAGGAGTCCAGAGATCCTAATNGG	555
Qy	120	YasnLysAlaAspGluGluInLysArgGlnValGlyArgGluGlnGlyInLeuAl	140

Db	556 GAATAAGCGCTGATGAACAGCAGAAACGGCCAGGTGGGAGAGAGAGAGAGAGAG-CTGGC 614
Qy	140 aLysGluTyrGlyMetAspPheTyrGluThr-SerAlaCysThrAsnLeuAsnIleLysG 160

[illegible]

Email: c9ap05-remail.nih.gov
Tissue Procurement: AFCC/DCTR/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI717 row: 1 column: 06
High quality sequence stop: 629.

```

FEATURES
source
1. .814
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4860437"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming, directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library. !"
223 a 252 q 133 t
BASE COUNT

```

BASE COUNT	223 a	206 c	252 q	133 t
BASE COUNT	223 a	206 c	252 q	133 t

ORIGIN

Alignment Scores:

Assignment Score:	8.53e-92	Length:	814
Pred. No.:	799.00	Matches:	161
Score:	96.43%	Conservative:	1
Percent Similarity:	96.43%	Mismatches:	3
Best Local Similarity:	95.83%	Indels:	3
Query Match:	72.31%	Gaps:	0
DB:	12		

US-09-817-198A-2 (1-212) X BG762967 (1-814)

QY 48 LysThrIleGluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGln 67

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAMI1807 row: e column: 13
High quality sequence stop: 777.
Location/Qualifiers

FEATURES

1. 911
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5318388"
/clone_lib="NCI CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 217 a 253 c 244 g 195 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 6.8e-89 Length: 911
Score: 777.00 Matches: 157
Percent Similarity: 98.16% Conservative: 3
Best Local Similarity: 96.32% Mismatches: 3
Query Match: 70.32% Indels: 2
DB: 13 Gaps: 0

US-09-817-198a-2 (1-212) x BF1649317 (1-911)

QY 50 IleGluValAspGlyIleLysValArgIleGlnIleTyrAspThrAlaGlyGlnGluArg 69
Db 1 ATCGATGTAGAGCGCATCAAGTGAATACAGATTGGGACACACAGCGAG-GAGAGG 59
QY 70 TyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlyIlePheLeuValTyrAsp 89
Db 60 TACCAGACTATCACAAAGCAGTACTATCGCGAGGCCA-GGAATATTTTAGTCTACGAC 118
QY 90 IleSerSerGluArgSerTyrGlnHisIleMetLysTyrIleValSerAspValAspGluTyr 109
Db 119 ATTAGCAGTGAGCGCTCTATCAGCATATCATGAAGTGGTGCAGTGGATGAGTAC 178
QY 110 AlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArg 129
Db 179 GCTCCAGAGAGTCCAGAGATCTTAATGGGAATAAGCTGTATGAGAGCAAGACG 238
QY 130 GlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGlu 149
Db 239 CAGGTGGGAGAGAGCAGGGGAGCAGCTGGCTAAGGAGTACGGCATCTTCTACGAA 298
QY 150 ThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuVal 169
Db 299 ACAAGTGCCTGCACCAACCTTAATATAAGAGTCTTCCACCTGCTGACGAGCTGGTG 358
QY 170 LeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAla 189
Db 359 CTCGAGGCCCCAGAGAAAGAGCTGGATGTCTCCGACACAGTCCGCAACAGCTCGCA 418
QY 190 LeuAlaGluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSertLysThr 209
Db 419 CTGGCGGAGCTGGAGGAGCAAGGCAACCTGAGGGGCCAGCAAACTCTTCAAGACC 478
QY 210 CysTrpCys 212
Db 479 TGCTGGTGC 487

RESULT 14
BF150465
LOCUS BF150465 618 bp mRNA linear EST 29-DEC-2000

DEFINITION

uy86b05.y1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3666417 5' similar to SW:RB15_RAT P35289 RAS-RELATED PROTEIN RAB-15. ;, mRNA sequence.

ACCESSION BF150465
VERSION BF150465.1 GI:11031860
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 618)

REFERENCE

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/iresources.shtml

MGI:1427185

Seq primer: -40RP from Gibco
High quality sequence stop: 419.

FEATURES

source

1. 618
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3666417"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"

BASE COUNT 145 a 166 c 194 g 111 t 2 others
ORIGIN

Alignment Scores:

Pred. No.: 1.83e-86 Length: 618
Score: 756.00 Matches: 146
Percent Similarity: 94.81% Conservative: 0
Best Local Similarity: 94.81% Mismatches: 8
Query Match: 68.42% Indels: 0
DB: 12 Gaps: 0

US-09-817-198a-2 (1-212) x BF150465 (1-618)

QY 1 MetaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
Db 157 ATGGCGAAACAGTACGATGCTGTTCGGCTACTCTGATCGGGACTCCGGGGTTGC 216
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeSerThr 40
Db 217 AAGACATGCTGCTGTGCGCTTACCAGCAACGAGTCTCCTCCGATATCTCCACC 276
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 277 ATCGGTGTTGACTTTAAGATGAAGACTATCGAAGTAGAGCGCATCAAGTGAATACAG 336
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
Db 337 ATTTGGACACAGCAGGGCAGGAGGTACAGACTATCACAAAGCAGTACTATCGCGCA 396
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100

```
Db 397 GCCCAGGGAATATTNTAGTCTAGCACATTAGCAGTGAGCGCTCCTATCAGCATATCATG 456
Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 457 AAGTGGGTCACTGACGTGGATGAGTACCTCTCANGAGGAGTCCAGAAGATCCTAATTGGG 516
Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140
Db 517 AATAATGCTGATGAAGAGCAACACCGAGCTGGGGAGAGACAGGTCACACGCTGGC 576
Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThr 154
Db 577 TTAGTAGTCGATGGAGCTTCTAGCAACAAAGTGCTGCACC 618

RESULT 15
LOCUS BM462147
DEFINITION AGENCOURT_6424626 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5491242
5', mRNA sequence.
ACCESSION BM462147
VERSION BM462147.1 GI:18511187
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1767)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12111 row: g column: 19
High quality sequence start: 89
High quality sequence stop: 574.
Location/Qualifiers
1..1767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5491242"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 318 a 704 c 411 g 320 t 14 others
ORIGIN

Alignment Scores:
Pred. No.: 8.62e-83 Length: 1767
Score: 733.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.33% Indels: 0
DB: 13 Gaps: 0

US-09-817-198A-2 (1-212) x BM462147 (1-1767)

Qy 72 ThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuValTyrAspIleSer 91
Db 96 ACCATCAAAAGCAGTACTATCGCGGGCCCGAGGATATTTTGGTCTATGACATTAGC 155
Qy 92 SerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyrAlaPro 111
```

```
Db 156 AGCGAGCGCTCTTACCAGCACATCATGAAGTGGGTCACTGACGTGGATGAGTACGCACCA 215
Qy 112 GluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArgGlnVal 131
Db 216 GAAGCGCTCCAGAAGATCCTTATTGGGAATAAGGCTGATGAGGAGCAGAAAACGGCAGGTG 275
Qy 132 GlyArgGluGlnGlyGlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGluThrSer 151
Db 276 GGAAGAGAGCAAGGCGACAGCTGCCGAGGAGTATGTCATGGACTTCTATGAAACAAGT 335
Qy 152 AlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db 336 GCCTGCACCAACCTCAACATTAAAGAGTCATTACGCGCTCTGACAGAGCTGGTGTCTGCAG 395
Qy 172 AlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAla 191
Db 396 GCCCATAGGAAGGAGCTGGAAGGCTCCGGATGCGTGCAGCAATGAGTTGGCACTGGCA 455
Qy 192 GluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211
Db 456 GAGCTGGAGAGGAGGAGGAGGCAACCCGAGGCGCCGCAACTCTTCGAAAAACCTGCTGG 515
Qy 212 Cys 212
Db 516 TGC 518
```

Search completed: November 17, 2002, 22:05:03
Job time : 1962 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	798	72.2	168	11	Q91YW0	Q91yw0 mus musculus
2	557	50.4	211	5	Q9TYS2	Q9Tys2 caenorhabdi
3	546.5	49.5	204	5	O15971	O15971 drosophila
4	541	49.0	207	11	Q8VCF6	Q8vcf6 mus musculus
5	528.5	47.8	207	5	O18338	O18338 drosophila
6	526.5	47.6	200	4	Q9H0P3	Q9H0t3 homo sapien
7	526	47.6	202	11	Q9DD03	Q9hd03 mus musculu
8	521	47.1	206	3	Q9HEP4	Q9het4 aspergillus
9	519	47.0	216	10	O24466	O24466 arabidopsis
10	518	46.9	214	10	Q40218	Q40218 lotus japon
11	515.5	46.7	216	10	Q40215	Q40215 lotus japon
12	515	46.6	216	10	Q9FJF1	Q9fjf1 arabidopsis
13	514	46.5	201	5	Q94148	Q94148 caenorhabdi
14	513.5	46.5	215	10	Q40219	Q40219 lotus japon
15	513	46.4	216	10	O8VMF9	Q8vwf9 nicotiana t
16	512	46.3	216	10	Q40177	Q40177 lycopersico

Db	121	NKADEQKRVGREGQQQLAKEYGMDFYETSAC	153
RESULT 2			
Q9TVS2			
ID	Q9TVS2	PRELIMINARY;	PRT; 211 AA.
AC	Q9TVS2;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	C. elegans RAB-8 protein (corresponding sequence D1037.4).		
GN	RAB-8.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Peloderinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		
RX	MEDLINE=99069613; PubMed=9851916;		
RA	Waterston R.;		
RT	"Genome sequence of the nematode C. elegans: a platform for		
RT	investigating biology. The C. elegans Sequencing Consortium."		
RL	Science 282:2012-2018(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		
RA	Ledwith J., Biewald T.;		
RT	"The sequence of C. elegans cosmid D1037.4";		
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		
RA	Waterston R.;		
RT	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AF106592; AAK21367.2;		
DR	SEQUENCE 211 AA; 24022 MW; B7609A91B6082DA2 CRC64;		
Qy			
Query Match	50.4%; Score 557; DB 5; Length 211;		
Best Local Similarity	55.0%; Pred. No. 6.9e-39;		
Matches 104; Conservative	42; Mismatches 39; Indels 4; Gaps 1;		
Qy	1 MAKQYDLFRLLIGDSGVGKTCCLCRFTDNEPHSSHISTIGVDFKMTIEVDGKIVRIQ	60	
Db	1 MAKYDYLFKLLIGDSGVGKTCVLFERFSDSFFNSFISTIGIDFKIRTIELDGKKIKLQ	60	
Qy	61 IWDTAGQRYOTITKQYRRAGQIFLYVDISSERSYQHMKWSDVDVEYAPGVQKILIG	120	
Db	61 IWDTAGQRFRTITATYRGAMGILYVDITNERSFENIKWIRNIEHAASDVVERMIIG	120	
Qy	121 NKADEQKRVGREGQQQLAKEYGMDFYETSACNLNLIKESFTRLTELVLQAHKKEGL	180	
Db	121 NKCDTEERVSRRGEGQLAIEYGTGKFLETSAKANLNIDEAFFTLARDI-----KSKMEQN	176	
Qy	181 RMRASNELA	189	
Db	177 EMRAGGSVS	185	
RESULT 3			
ID	O15971	PRELIMINARY;	PRT; 204 AA.
AC	O15971;		
DT	01-JAN-1998 (TrEMBLrel. 05, Created)		
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)		
DE	RAB10 protein (LD39986P).		
GN	RAB10 OR DRAB10 OR CG17060.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		


```
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: AE003515; AAF49101.1; -.
DR EMBL: D84347; BAA21711.1; -.
DR HSSP: P05713; 3RAB.
DR FlyBase: FBgn0015796; Rab8.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001230; Prenyl_site.
DR InterPro: IPR001806; Ras_trnsmg.
DR InterPro: IPR002078; Sig54_interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 207 AA; 23691 MW; BE9DE812C77DCF09 CRC64;

Query Match 47.8%; Score 528.5; DB 5; Length 207;
Best Local Similarity 53.8%; Pred. No. 1.6e-36;
Matches 100; Conservative 41; Mismatches 40; Indels 5; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTIYVDGKIRIQ 60
Db 1 MAKTYDYLKLLIGDSGVGKTCILFRSEDAFNFTTISTIGIDFKIRTELDNKKIKLQ 60

QY 61 IWDTAGQERYQTITKQYRRAOGIFLVYDISSERSYOHIMKWVSDVDEYAPGEGVKILIG 120
Db 61 IWDTAGQERFTITAYRGANGIMLYDITQKSFENIKRWIENENASADVCKMLLG 120

QY 121 NKADEQKRVGRGQOQLAKEYGMDFYTSACTNLNLIKESFTRLTVELVQAHKKELEGL 180
Db 121 NKCLTDKRVQSKRGQLAEYGIKFMETSAKASINVEEAF-----LTLASDIKAKTEK 175

QY 181 RMRASN 186
Db 176 RMEANN 181

RESULT 6
ID Q9HOT3 PRELIMINARY; PRT; 200 AA.
AC Q9HOT3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 22.5 kDa protein.
GN DKFZF564L1962.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21154917; PubMed=11230166;
RA Wilmann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansoerg W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL Genome Res. 11:422-435(2001).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: AL136650; CAB6585.1; -.
DR HSSP: P05713; 3RAB.
```

```
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsmg.
DR InterPro: IPR002078; Sig54_interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Hypothetical protein; Lipoprotein.
SQ SEQUENCE 200 AA; 22469 MW; 7F01DB88E46EE3EA CRC64;

Query Match 47.6%; Score 526.5; DB 4; Length 200;
Best Local Similarity 55.8%; Pred. No. 2.3e-36;
Matches 96; Conservative 45; Mismatches 30; Indels 1; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMKTIYVDGKIRI 59
Db 1 MAKTYDYLKLLIGDSGVGKTCVLFRRSDDAFNTTFTSTIGIDFKIKTVELOGKKIKL 60

QY 60 QIWDTAGQERYQTITKQYRRAOGIFLVYDISSERSYOHIMKWVSDVDEYAPGEGVKILI 119
Db 61 QIWDTAGQERFTITAYRGANGIMLYDITNGKSEFNISKWLRNIDEHANEDVERMLL 120

QY 120 GNKADEEQKRVGRGQOQLAKEYGMDFYTSACTNLNLIKESFTRLTVELVLQ 171
Db 121 GNKCDMDKRVKPGKGQIAREHGIRFFETSAKANINIEKAFLLTAEILR 172

RESULT 7
ID Q9DD03 PRELIMINARY; PRT; 202 AA.
AC Q9DD03;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 0610007N03RIK protein (RIKEN cDNA 0610007N03 gene).
GN 0610007N03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wuyishaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaishizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
SEQUENCE FROM N.A.
RC Strausberg R.;
RX Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: AK022303; BAB22000.1; -.
DR EMBL: BC027214; AAB27214.1; -.

```



```
DR HSSP; P05713; 3RAB.
DR MGD; MGI:1915578; 0610007N03RIK.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFS; TIGR00231; small_gtp; 1.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 202 AA; 22770 MW; 5DF599432E228AC0 CRC64;

Query Match 47.6%; Score 526; DB 11; Length 202;
Best Local Similarity 54.1%; Pred. No. 2.5e-36;
Matches 92; Conservative 45; Mismatches 33; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDFKMTIEVDGIKVRIO 60
DB 1 MAKADHLFLKLLIGDSGVGKTCCLIRFAEDNFNFSTYISTIGIDPKITVDIEGRIKIQ 60

QY 61 IWDTAGQERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIG 120
DB 61 VMDTAGQERFKTITAYYRGAMGILLVYDITDEKSFENIQNMKSIKENASAGVERLLG 120

QY 121 NKADEQKRVGREGOQOLAKKEYGMDFYETSACTNLNLIKESFTRLTELVL 170
DB 121 NKCDMEARKVQREQAKELAREHRIRFFETSASKSVNVDFAFSSLDAIL 170

RESULT 8
Q9HET4 PRELIMINARY; PRT; 206 AA.
AC Q9HET4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Secretion related GTPase, (SrgA).
GN SRGA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N402;
RA Konetschny C., Ram A.F., Montijn R., Kubicek C.P.,
RA Van Den Hondel C.A.;
RT "Identification and characterisation of a family of secretion related
small GTPase encoding genes from the filamentous fungus Aspergillus
niger: a putative SRC4 homologue is not essential for growth.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR HSSP; P05713; 3RAB.
DR EMBL; AJ278658; CAC17832.1; -.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFS; TIGR00231; small_gtp; 1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 206 AA; 22823 MW; 501916B795CF8C9C CRC64;

Query Match 47.1%; Score 521; DB 3; Length 206;
Best Local Similarity 60.6%; Pred. No. 6.8e-36;
Matches 97; Conservative 30; Mismatches 33; Indels 0; Gaps 0;
```

```
QY 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDFKMTIEVDGIKVRIO 62
DB 5 RNYDFLIKLLIGDSGVGKSCCLLRFSDFTPTTIGIDFKIRTIELDGKRVKLQIW 64

QY 63 DTAGQERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIGNK 122
DB 65 DTAGQERFTITAYYRGAMGILLVYDITDEKSFENIQNMKSIKENASAGVERLLG 124

QY 123 ADEQKRVGREGOQOLAKKEYGMDFYETSACTNLNLIKESF 162
DB 125 CDWEKRAVSTEQGOQLANELGIPPLEVSAKNNINIEKAF 164

RESULT 9
O24466 PRELIMINARY; PRT; 216 AA.
AC O24466;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE ATRAB8 (GTPase ATRAB8).
GN RAB8 OR F4P12.310.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Bischoff F., Palme K.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bloeker H., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,
RA Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; U82434; AAB65088.1; -.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFS; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 216 AA; 23939 MW; 1648E45B29D4EEB6 CRC64;

Query Match 47.0%; Score 519; DB 10; Length 216;
Best Local Similarity 51.9%; Pred. No. 1.1e-35;
Matches 98; Conservative 42; Mismatches 43; Indels 6; Gaps 2;

QY 5 YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDFKMTIEVDGIKVRIO 64
DB 12 YDLKLLIGDSGVGKSCCLLRFSDGFTTIGIDFKIRTIELDGKRIKIQIW 71

QY 65 AGQERYQTITKQYRRAGIFLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIGNK 124
DB 72 AQGERFTITAYYRGAMGILLVYDITDEKSFENIQNMKSIKENASAGVERLLG 131

QY 125 -BEQKRVGREGOQOLAKKEYGMDFYETSACTNLNLIKESFTRLTELVLQA----HRKELE 178
DB 132 MDESKRAVPSKSGQALADEYGMKFFETSAKTNLNVVEVFFSTAKDIKORLADTDARAEPQ 191
```

```
QY 179 GLRRASNE 187
DB 192 TIKINSDQ 200

RESULT 10
Q40218 PRELIMINARY; PRT; 214 AA.
AC Q40218;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE RAB8D.
GN RAB8D.

OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT NODULES;
RX MEDLINE=97231679; PubMed=9076991;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
binding proteins encoded by cDNAs from Lotus japonicus, and expression
of corresponding mRNAs in developing root nodules.";
RL Plant J. 11:237-250(1997).
CC -|- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z73944; CAA98172.1; -.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 216 AA; 23776 MW; 569926CCA8D1B489 CRC64;

Query Match 46.7%; Score 515.5; DB 10; Length 216;
Best Local Similarity 48.3%; Pred. No. 2.1e-35;
Matches 100; Conservative 43; Mismatches 57; Indels 7; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCILCRFTDNEFHSHSTIGVDFKMKTIQVGIKVIQIWD 64
DB 13 YDYLKLLIGDSGVGKSCLLRFSFGSFTTIFIGDKTIELDGKRLQIWD 72
QY 65 AGQERYQTITKQYRRAQGIFLYVDISSERSYQHIMKWSVDVDEYAPGVOKILGNKAD 124
DB 73 AGQERFTITAYRGAMGILLVYDVEASFNNIRNIRNIEQHASDNVKNKILVGNKAD 132
QY 125 -EOKRVGREGQOLAKYGMDFYETSACTNLNIKESFTRLTFLVLAHQHKEGLRMR 183
DB 133 MDESKRAVPTSKQALADEYGIKFETSAKTNLNVVEVFSIARDI----KQRLADTDSR 188
QY 184 ASNELALAELEEKGPEGPANSSKTC 210
DB 189 A--EPQTIQINQPDASASGGQAQKSC 213

RESULT 12
Q9FJF1 PRELIMINARY; PRT; 216 AA.
AC Q9FJF1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Rab-type small GTP-binding protein-like.
OC Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=99087489; PubMed=9872454;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
RT Sequence features of the regions of 1,013,767 bp covered by sixteen
RT physically assigned pl and TAC clones.";
RL DNA Res. 5:297-308(1998).
```

```

RESULT 13
Q94148
ID Q94148 PRELIMINARY; PRT; 201 AA.
AC Q94148;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE RAB10-like.
OS T23H2.5.
NC CAenorhabditis elegans.
NC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Wamsley P., Bradshaw H.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RW [2]
RS SEQUENCE OF 1-124 FROM N.A.
RA Nonet M.L., Staunton J.E., Kilgard M.P., Fergestad T., Jorgensen E.,
RA Hartweig E., Horvitz H.R., Meyer B.J.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; U80033; AAC48200.1; -.
DR EMBL; U68257; AAB16972.1; -.
DR HSSP; P05713; RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_Grnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.

```

Query Match	46.5%	Score 513.5;	DB 10;	Length 215;
Best Local Similarity	46.9%;	Pred. No. 3.1e-35;		
Matches 100;	Conservative 43;	Mismatches 53;	Indels 17;	Gaps 3;
QY	5	YDVLFRLLLGDSGVGKTCLLCRFTDNEFHSSHSHTSTGVDFFKMTIEVDGKIVRIQIWD	64	
		: : : : :		
Db	12	YDVLIKLLGDSGVGKSCLLRFSDGTSFTTFTTIGIDFKIRTIELDGKRKLIQIWD	71	
QY	65	AGERYOTIKQYRRQAQITFLVYDISERSYQHIMKWSDVDYAEQVQKILIGNKAD	124	
		: : : : :		
Db	72	AGOREFTITATYYRGAMGILLVYDVTDSESNFKNIWNIHQHSDNVNKKILVGNKAD	131	

QY 125 -EEQKRVGREGQQQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQ-----AHRKEL 177
Db 132 MDESKRAVPTSKQALADEYGIKFFETSATNLNVEEVFFSIARDIKQRLADTDSKTEP 191
QY 178 EGLMRASNELALAELEEEGKPEGPANSKTC 210
Db 192 TGIKIN-----PQKGSAGEAAQKSACC 214

RESULT 15

OBVWF9 PRELIMINARY; PRT; 216 AA.
AC Q8VWF9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Ras-related protein RAB8-4 (Ras-related protein RAB8-2).
GN RAB8-4 OR RAB8-2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRIGHT YELLOW 2;
RA Torimoto N., Shimada K., Ito K., Yamamoto K.;
RT "Characterization of Rab8 from tobacco BY-2 cell."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB079023; BAB84325.1; -.
DR EMBL; AB079021; BAB84323.1; -.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfrmng.
DR InterPro: IPR002078; Sig54_interact.
DR Pfam: PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
SQ SEQUENCE 216 AA; 23971 MW; D5E87FABD0D49E13 CRC64;

Query Match 46.4%; Score 513; DB 10; Length 216;
Best Local Similarity 46.7%; Pred. NO. 3.4e-35;
Matches 99; Conservative 44; Mismatches 53; Indels 16; Gaps 3;
QY 5 YDVLFRLLIGDSGVGKTCILCRFTDNEFHSHISTIGVDFKMKTIIEVDGIKVIQIWD 64
Db 12 YDVLKILLIGDSGVGKSCILLRFSFGSFTTIGIDFKIRTIELDGKRIKLIQIWD 71
QY 65 AGQERYQTITKQYRRAQGIPLVYDISSERSYOHIMKWSDVDVEYAPGVQKILIGNKAD 124
Db 72 AGQERFTITTAIRGAMGILLVYDVTDESSFNIRNIRNIEQHASDNVKNKILVGNKAD 131
QY 125 -EEQKRVGREGQQQLAKYGMDFYETSACTNLNLIKESFTRLTELVLQ-----HRKELE 178
Db 132 MDESKRAVPTSKQALADEYGIKFFETSATNLNVEEVFFSIARDIKQRLSEDSKTEPQ 191
QY 179 GLRMRASNELALAELEEEGKPEGPANSKTC 210
Db 192 AIRNQSDQAGTS-----GQAQKSSC 213

Search completed: November 17, 2002, 20:40:55
Job time : 83 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 18:15:09 ; Search time 25 Seconds

(without alignments)
351.719 Million cell updates/sec

Title: US-09-817-198A-2

Perfect score: 1105

Sequence: 1 MAKQYDLVRLLLIGDSGVG.....LEEERKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1077	97.5	212	1 RB15_RAT	P35289 rattus norv
2	545	49.3	207	1 RAB8_HUMAN	P24407 homo sapien
3	539.5	48.8	210	1 RAB8_DISOM	P22128 discopoyge o
4	532	48.1	203	1 RB13_HUMAN	P51153 homo sapien
5	531.5	48.1	200	1 RAO1_DISOM	P22127 discopoyge o
6	530.5	48.0	200	1 RB10_CANFA	P24409 canis famli
7	530.5	48.0	200	1 RB10_HUMAN	O88386 homo sapien
8	528.5	47.8	207	1 RB8B_HUMAN	Q92930 homo sapien
9	528.5	47.8	207	1 RB8B_RAT	P70550 rattus norv
10	526	47.6	206	1 RAB8_MOUSE	P55258 mus musculu
11	510	46.2	200	1 YPT2_SCHPO	P17609 schizosach
12	509	46.1	215	1 RAB1_BETVU	Q39433 beta vulgar
13	508	46.0	216	1 ARA3_ARATH	P28186 arabidopsis
14	505.5	45.7	217	1 YPT2_VOLCA	P36861 volvox cart
15	502.5	45.5	200	1 RB10_RAT	P35281 rattus norv
16	502	45.4	203	1 RYL1_YARLI	P41924 yarrowia li
17	483.5	43.8	208	1 SAS1_DICDI	P20790 dictyosteli
18	483	43.7	208	1 YPT1_MAIZE	P16976 zea mays (m
19	480	43.4	202	1 RIC1_ORISA	P40392 oryza sativ
20	479	43.3	203	1 SAS2_DICDI	P20791 dictyosteli
21	477.5	43.2	201	1 RB35_HUMAN	Q15286 homo sapien
22	477.5	43.2	203	1 YPT1_CHLRE	Q39571 chlamydomon
23	477	43.2	201	1 YPT1_PHYIN	Q01890 phytophor
24	476	43.1	258	1 ARA5_ARATH	P28188 arabidopsis
25	475.5	43.0	203	1 YPT1_VOLCA	P31584 volvox cart
26	475	43.0	203	1 YPT1_NEUCR	P33723 neurospora
27	473	42.8	210	1 SEC4_CANAL	Q14462 candida alb
28	470.5	42.6	203	1 YPT2_MAIZE	O05737 zea mays (m
29	470.5	42.6	215	1 SEC4_YEAST	P07560 saccharomyc
30	469	42.4	201	1 RB1B_RAT	P10536 rattus norv
31	469	42.4	205	1 RAB1_LYMST	Q05974 lymanaea sta
32	468	42.4	205	1 RB1A_RAT	P05711 rattus norv
33	467	42.3	203	1 YPT1_SCHPO	P11620 schizosach

ALIGNMENTS

RESULT 1

ID	RB15_RAT	STANDARD;	PRT;	212 AA.
AC	P35289;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Ras-related protein Rab-15.			
GN	RAB15.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;			
RX	MEDLINE=92210533; PubMed=1313420;			
RA	Elferink L.A., Anzai K., Scheller R.H.;			
RT	"Rab15, a novel low molecular weight GTP-binding protein specifically expressed in rat brain."			
RL	J. Biol. Chem. 267:5768-5775(1992).			
RN	[2]			
RP	ERRATUM.			
RX	MEDLINE=93054572; PubMed=1429617;			
RA	Elferink L.A., Anzai K., Scheller R.H.;			
RL	J. Biol. Chem. 267:22693-22693(1992).			
CC	-!- FUNCTION: MAY ACT IN CONCERT WITH RAB3A IN REGULATING ASPECTS OF SYNAPTIC VESICLE MEMBRANE FLOW WITHIN THE NERVE TERMINAL.			
CC	-!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN NEURAL TISSUES.			
CC	-!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M83679; AAA41995.1; -			
DR	PIR; F42148; F42148.			
DR	HSSP; P05713; 3RAB.			
DR	InterPro; IPR003579; GTPase_Rab.			
DR	InterPro; IPR001806; Ras_trnsfrmg.			
DR	InterPro; IPR005325; Small_GTP.			
DR	Pfam; PF00071; ras; 1			
DR	PRINTS; PR00449; RASTRNSFRMG.			
DR	SMART; SM00175; RAB; 1.			
DR	TIGRFAMs; TIGR00231; small_GTP; 1.			
KW	GTP-binding; Lipoprotein; Prenylation; Protein transport.			
FT	NP_BIND 15 22 GTP (BY SIMILARITY).			
FT	NP_BIND 63 67 GTP (BY SIMILARITY).			
FT	NP_BIND 121 124 GTP (BY SIMILARITY).			
FT	LIPID 210 210 GERANYL-GERANYL (BY SIMILARITY).			
FT	LIPID 212 212 GERANYL-GERANYL (BY SIMILARITY).			
SQ	SEQUENCE 212 AA; 24283 MW; 04817DDA66CADE12 CRC64;			

P11476 homo sapien
P22125 discopoyge o
P01123 saccharomyc
Q63482 rattus norv
Q9ulw5 homo sapien
P51156 rattus norv
P25228 drosophila
O95716 homo sapien
P11023 bos taurus
Q63942 rattus norv
P05713 mus musculu
P20336 homo sapien

```
Query Match          97.5%; Score 1077; DB 1; Length 212;
Best Local Similarity 97.6%; Pred. No. 6.9e-80;
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLRCFTDNEFHSHSHSTIGVDFKMKTIYVDGKIKVRIQ 60
Db 1 MAKQYDVLFRLLIGDSGVGKTCCLRCFTDNEFHSHSHSTIGVDFKMKTIYVDGKIKVRIQ 60

QY 61 IWDTAGQERYOTITKQYRRAGQIFLYVDISSERSYQHIMKMWSDVDYVAPGVQKILIG 120
Db 61 IWDTAGQERYOTITKQYRRAGQIFLYVDISSERSYQHIMKMWSDVDYVAPGVQKILIG 120

QY 121 NKADEQKRVQREGQOQLAKKEYGMDFYETSACTNINIKESFTRRLTELVLQAHKKEGLE 180
Db 121 NKADEQKRVQREGQOQLAKKEYGMDFYETSACTNINIKESFTRRLTELVLQAHKKEGLE 180

QY 181 RMRASNE 187
Db 181 SPOGSNQ 187

RESULT 2
RAB8_HUMAN STANDARD; PRT; 207 AA.
AC P24407;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein rab-8 (Oncogene c-mel).
GN RAB8 OR MEL.
OS Homo sapiens (Human), and
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 9615;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=94124602; PubMed=8294494;
RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,
RA Tavetian A., Louvard D.;
RT "A small rab GTPase is distributed in cytoplasmic vesicles in non
RT polarized cells but colocalizes with the tight junction marker ZO-1
RT in polarized epithelial cells.";
RL J. Cell Biol. 124:101-115(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=91360267; PubMed=1886711;
RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,
RA Johnson K.J.;
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related
RT genes.";
RL Oncogene 6:1347-1351(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=C.familiaris; STRAIN=Cocker spaniel;
RX MEDLINE=91061765; PubMed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of Yp1/SEC4-related cDNAs from an epithelial cell
RT line.";
RL Mol. Cell. Biol. 10:6578-6585(1990).
CC -I- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE.
CC -I- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56741; CAA40065.1; -.
DR EMBL; S53268; AAB19681.1; -.
DR EMBL; BC002977; AAH02977.1; -.
DR EMBL; X56385; CAB56776.1; -.
DR PIR; B36364; B36364.
DR PIR; B49647; B49647.
DR HSSP; P05713; 3RAB.
DR Genew; HGNC:7007; MEL.
DR MIM; 165040; -.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR001806; Ras_trnsmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW Proto-oncogene.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
FT CONFLICT 177 183 LEGNSPQ -> WKATAP (IN REF. 2).
SQ SEQUENCE 207 AA; 23668 MW; AA52DBF54A2CD056 CRC64;

Query Match          49.3%; Score 545; DB 1; Length 207;
Best Local Similarity 52.9%; Pred. No. 4.2e-37;
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLRCFTDNEFHSHSHSTIGVDFKMKTIYVDGKIKVRIQ 60
Db 1 MAKQYDVLFRLLIGDSGVGKTCCLRCFTDNEFHSHSHSTIGVDFKMKTIYVDGKIKVRIQ 60

QY 61 IWDTAGQERYOTITKQYRRAGQIFLYVDISSERSYQHIMKMWSDVDYVAPGVQKILIG 120
Db 61 IWDTAGQERYOTITKQYRRAGQIFLYVDISSERSYQHIMKMWSDVDYVAPGVQKILIG 120

QY 121 NKADEQKRVQREGQOQLAKKEYGMDFYETSACTNINIKESFTRRLTELVLQAHKKEGLE 180
Db 121 NKCDVNDKRVQSKERGEKALDYGIKFMETSAKANINVENAFPTLARDIKAKMDKKLEGN 180

QY 181 RMRASNE 187
Db 181 SPOGSNQ 187

RESULT 3
RAB8_DISOM STANDARD; PRT; 210 AA.
AC P22128;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein Rab-8 (ORA2).
OS Discoprya ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hypnosqualea; Pristiogorae; Batoidea;
OC Torpediniformes; Narcinoidei; Narcinidae; Discopyge.
OX NCBI_TaxID=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Electric lobe;
RX MEDLINE=91115900; PubMed=1899244;
RA Ngsee J.K., Elferink L.A., Scheller R.H.;
```

CC TISSUE SPECIFICITY: VARIETY OF EPITHELIA, INCLUDING INTESTINE,
CC KIDNEY, LIVER, AND OF ENDOTHELIAL CELLS.
CC
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC	---	
DR	EMBL: X75593; CAA53266.1; --	
DR	EMBL: BC000799; AAH00799.1; --	
DR	HSSP: P01112; 1pLL	
DR	Genew: HGNC:9762; RAB13.	
DR	MIM: 602672; --	
DR	InterPro: IPR003579; GTPase_Rab.	
DR	InterPro: IPR001230; Prenyl_site.	

DR InterPro: IPR001806; Ras_trnsfrmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR

[illegible]

```

Db      1  MAKAYDHLFKLLILGDSGVCKTCLIRPAEDNFNTVISTIGIDFKITVDIEGKKIKLQ  60
QY      61  IWDTAGQERYOTITKQYRRAGQIFLYVDISSERSYQHIMKWSVDVDEYAPEGVQKILIG  120
Db      61  VMDTAGQERFKTITAYYRGAMGILYVDITDEKSFENIQNMKSIKENASAGVERLLIG  120
QY      121  NKADEQKQVGRBQGOQLAKCYGMDPYETSACTNLNKESFRLTTLVTLQAHKEBGL  180
Db      121  NKCDMEAKRRVQKEQADKLAREHGIRFETFSYAKSSMMVNDFAFSSLARDIL-----LKSG  174

```

```

QY 181 RMRASNELALAELEEEKKGPGPANSKTC 210
      || |
      || |
Db 175 GRRSGN-----GNKP--PSTDUKTC 192

RESULT 5
RAOL_DISOM
ID RAOL_DISOM STANDARD; PRT; 200 AA.
AC P22127;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)

```

DE Ras-related protein ORAL.
OS Discozyme ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hynostomalea; Pristiorajae; Batoidae;
OC Torpediniformes; Narcinoidae; Narcinidae; Discozyme.
OX NCBI_TaxID=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Electric lobe;
RX MEDLINE=91115900; PubMed=1899244;
RA Ngse J.K., Elferink L.A., Scheller R.H.;
RT "A family of ras-like GTP-binding proteins expressed in electromotor
neurons."
RL J. Biol. Chem. 266:2675-2680(1991).
CC -!- SIMILARITY: Ras-RELATED PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M38390; AAA49230.1; -.
DR PIR; A38625; A38625.
DR HSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_GTP; 1.
KW GTP-binding; 16 23 GTP (BY SIMILARITY).
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 122 125 GTP (BY SIMILARITY).
FT EFFECTOR REGION (BY SIMILARITY).
FT LIPID 38 46 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 200 AA; 22623 MW; 41D38E3D760519C5 CRC64;

Query Match 48.1%; Score 531.5; DB 1; Length 200;
Best Local Similarity 56.4%; Pred No. 4.9e-36;
Matches 97; Conservative 44; Mismatches 30; Indels 1; Gaps 1;

Qy 1 MAKQ-YDVLFRLLIGDSVGKTCVLRFDSESHSTIGVDFKMKTEVDGKIKV 59
Db 1 MAKTYDVLFRLLIGDSVGKTCVLRFDSESHSTIGVDFKMKTEVDGKIKV 60
Qy 60 QIWDTAGQERYQTITKQYRRAQIGFLVYDISSERSYQHIMKWSDYDEVAPEGVOKILI 119
Db 1 MAKTYDVLFRLLIGDSVGKTCVLRFDSESHSTIGVDFKMKTEVDGKIKV 60
Qy 60 QIWDTAGQERYQTITKQYRRAQIGFLVYDISSERSYQHIMKWSDYDEVAPEGVOKILI 119
Db 61 QIWDTAGQERYQTITKQYRRAQIGFLVYDISSERSYQHIMKWSDYDEVAPEGVOKILI 120
Qy 120 GNKADDEKQVREGQOQLAKEYGMDFYETSACTNINIKESFRLTELVLQ 171
Db 121 GNKCDMDKRVLPKSGEQIAREHAIFFETSAKINIEKAFLEADILQ 172
Qy 120 GNKADDEKQVREGQOQLAKEYGMDFYETSACTNINIKESFRLTELVLQ 171
Db 121 GNKCDMDKRVLPKSGEQIAREHAIFFETSAKINIEKAFLEADILQ 172

RESULT 6
RB10_CANFA STANDARD; PRT; 200 AA.
AC P24409;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Cocker spaniel;

RX MEDLINE=91061765; PubMed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
line".
RL Mol. Cell. Biol. 10:6578-6585(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
NEUROTRANSMITTER RELEASE.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56387; CAA39798.1; -.
DR PIR; D36364; D36364.
DR HSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 122 125 GTP (BY SIMILARITY).
FT EFFECTOR REGION (BY SIMILARITY).
FT LIPID 38 46 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 200 AA; 22569 MW; 5D52B8E47D4362 CRC64;

Query Match 48.0%; Score 530.5; DB 1; Length 200;
Best Local Similarity 55.8%; Pred No. 5.9e-36;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;

Qy 1 MAKQ-YDVLFRLLIGDSVGKTCVLRFDSESHSTIGVDFKMKTEVDGKIKV 59
Db 1 MAKTYDVLFRLLIGDSVGKTCVLRFDSESHSTIGVDFKMKTEVDGKIKV 60
Qy 60 QIWDTAGQERYQTITKQYRRAQIGFLVYDISSERSYQHIMKWSDYDEVAPEGVOKILI 119
Db 61 QIWDTAGQERYQTITKQYRRAQIGFLVYDISSERSYQHIMKWSDYDEVAPEGVOKILI 120
Qy 120 GNKADDEKQVREGQOQLAKEYGMDFYETSACTNINIKESFRLTELVLQ 171
Db 121 GNKCDMDKRVLPKSGEQIAREHAIFFETSAKINIEKAFLEADILR 172

RESULT 7
RB10_HUMAN STANDARD; PRT; 200 AA.
AC Q88386; Q9D7X6;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Homo sapiens (Human), and
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606; 10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Pituitary;
RX MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,

RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
 axis and full-length cDNA cloning";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Wong K., Hong W., Tang B.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoi T., Raku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Yamanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A.
 RC SPECIES=Mouse;
 RA Zeng Q., Tan Y.H., Hong W.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=C57BL/6J; TISSUE=Stomach;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Arawaka T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
 CC NEUROTRANSMITTER RELEASE.
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF106681; AAD43034.1; -;
 DR EMBL; AF297660; AAG34413.1; -;
 DR EMBL; AB023223; BAB14474.1; -;
 DR EMBL; BC000896; AAH00896.1; -;

DR EMBL; AF035646; AAC29313.1; -;
 DR EMBL; AK008725; BAB25858.1; -;
 DR HSSP; P05713; 3RAB.
 DR Genew; HGNC:9759; RAB10.
 DR MGD; MGI:105066; RAB10.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsfrmg.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
 FT NP_BIND 16 23 GTP (BY SIMILARITY).
 FT NP_BIND 64 68 GTP (BY SIMILARITY).
 FT NP_BIND 122 125 GTP (BY SIMILARITY).
 FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
 FT CONFLICT 106 106 N -> H (IN REF. 6).
 SQ SEQUENCE 200 AA; 22541 MW; 7F02B8E8E46EE1E8 CRC64;
 Query Match 48.0%; Score 530.5; DB 1; Length 200;
 Best Local Similarity 55.8%; Pred. No. 5.9e-36;
 Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
 QY 1 MAKQ-YDVLFRLLIGDSGVGKTCILCRFTDFNEFHSHSTIGVDPKMTIEVDGIKVI 59
 DB 1 MAKYYDLLFKLLIGDSGVGKTCVLFKFSDDAFNTFTSTIGIDFKITVELQKKIKL 60
 QY 60 QIWDTAGQERYQYITTKQYVRRAGIFLVYDISSERSYQHIMKWSVDVYAPGVOKILI 119
 DB 61 QIWDTAGQEREHTITTSYYRGAMIMLVYDITNGKSFENISKWLRNIDSHANEDVERMLL 120
 QY 120 GNADEEQKRVGRGCGQQLAKEYGMDFTYETACTNMLNKESTFTRTELVLQ 171
 DB 121 GNKCDMDRRVYPKGGEGQIAREHGIRFFETSAKANINIEKAFLTLAEDILR 172
 RESULT 8
 ID RB8B_HUMAN STANDARD; PRT; 207 AA.
 AC Q92930; Q9P293;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab-8B.
 GN RAB8B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RA Seki N., Saito T.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 67-119 FROM N.A.
 RC TISSUE=Melanoma;
 RX MEDLINE=97182150; PubMed=9030196;
 RT Chen D., Guo J., Gahl W.A.;
 RT "RAB GTPases expressed in human melanoma cells";
 RL Biochim. Biophys. Acta 1355:1-6(1997).
 CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
 CC NEUROTRANSMITTER RELEASE (BY SIMILARITY).
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB038995; BA92249.1; -
 DR EMBL; BC020654; AAH2654.1; -
 DR EMBL; U66624; AAC51199.1; -
 DR HSSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001230; Prenyl_site.
 DR InterPro; IPR001806; Ras_trnsmng.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 KW GTP-binding; Prenylation; Lipoprotein; Protein transport.
 FT NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 63 67 GTP (BY SIMILARITY).
 FT NP_BIND 121 124 GTP (BY SIMILARITY).
 FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 207 AA; 23584 MW; 5960993C0F87F944 CRC64;

Query Match 47.8%; Score 528.5; DB 1; Length 207;
 Best Local Similarity 54.0%; Pred. No. 8.9e-36;
 Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVKTCCLCRFTDNEPHSHSHISTIGVDFKMKTEVDGKVRQ 60
 DB 1 MAKTYDYLFRLLIGDSGVKTCCLCRFTDNEPHSHSHISTIGVDFKMKTEVDGKVRQ 60
 QY 61 IWDTAGERYQTITKQYRRAQGFILVYDISSERSYQHIMKWSVDVDEYAPEGVQKILIG 120
 DB 61 IWDTAGERFRTITTAIRGAMGIMLVYDITNEKSFNKNWIRNIEEHASSDVERMILG 120
 QY 121 NKADEEKQVGRGQOOLAKEYGMDFYETSACTNLNIKES-PTRLTELVLQHRK 175
 DB 121 NKCDMNDKQVSKERGEKLAIDYGIKLETSKASSNVEEAFFTLARDIMTKLNRK 176

RESULT 9
 RB8B_RAT STANDARD; PRT; 207 AA.
 AC P70550;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Ras-related protein Rab-8B.
 GN RB8B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96393028; PubMed=8799816;
 RA Armstrong J., Thompson N., Squire J.H., Smith J., Hayes B., Solari R.;
 RT "Identification of a novel member of the Rab8 family from the rat
 RT basophilic leukaemia cell line, RBL.2H3.";
 RL J. Cell Sci. 109:1265-1274(1996).
 CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
 CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS OF EXPRESSION IN THE SPLEEN,
 CC TESTIS AND BRAIN.
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U53475; AAA99782.1; -
 DR HSSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001230; Prenyl_site.
 DR InterPro; IPR001806; Ras_trnsmng.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 KW GTP-binding; Prenylation; Lipoprotein; Protein transport.
 FT NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 63 67 GTP (BY SIMILARITY).
 FT NP_BIND 121 124 GTP (BY SIMILARITY).
 FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 207 AA; 23603 MW; 4A41AB26BF9DCAF4 CRC64;

Query Match 47.8%; Score 528.5; DB 1; Length 207;
 Best Local Similarity 54.0%; Pred. No. 8.9e-36;
 Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVKTCCLCRFTDNEPHSHSHISTIGVDFKMKTEVDGKVRQ 60
 DB 1 MAKTYDYLFRLLIGDSGVKTCCLCRFTDNEPHSHSHISTIGVDFKMKTEVDGKVRQ 60
 QY 61 IWDTAGERYQTITKQYRRAQGFILVYDISSERSYQHIMKWSVDVDEYAPEGVQKILIG 120
 DB 61 IWDTAGERFRTITTAIRGAMGIMLVYDITNEKSFNKNWIRNIEEHASSDVERMILG 120
 QY 121 NKADEEKQVGRGQOOLAKEYGMDFYETSACTNLNIKES-PTRLTELVLQHRK 175
 DB 121 NKCDMNDKQVSKERGEKLAIDYGIKLETSKASSNVEEAFFTLARDIMTKLNRK 176

RESULT 10
 RB8B_MOUSE STANDARD; PRT; 206 AA.
 AC P52558;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Ras-related protein Rab-8 (Oncogene c-mel).
 GN RB8 OR MEL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91360267; PubMed=1886711;
 RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,
 RA Johnson K.J.;
 RT "The MEL gene: a new member of the RAB/YPT class of RAS-related
 RT genes.";
 RL Oncogene 6:1347-1351(1991).
 CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
 CC NEUROTRANSMITTER RELEASE.
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC -----
DR EMBL; S53270; AAB19682.1; -.
DR HSSP; P05713; 3RAB.
DR MGD; MGI:96960; Mel.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_gtp; 1.
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW Proto-oncogene.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 206 AA; 23557 MW; 49D832725D662942 CRC64;

Query Match 47.6%; Score 526; DB 1; Length 206;
Best Local Similarity 56.4%; Pred. No. 1.4e-35;
Matches 93; Conservative 41; Mismatches 31; Indels 0; Gaps 0;

QY 1 MAKQYDLVFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVRIQ 60
Db 1 MAKTYDLFKLLIGDSGVGKTCVLFRESDAFNSTFTIGIDFKIRTIELDGKRIKIQ 60

QY 61 IWDTAGQERYQTITKQYRRAQIGFLVYDISSERYOHIMKMWSDVDEYAPGVOKILIG 120
Db 61 IWDTAGQERFRTITPAYRGAMGIMLVYDITNEKSFNIRNIRNIEHASADVEKMILG 120

QY 121 NKADDEQKRVGREGOQOLAKRYGMDVETSACTNLNIKESFTRL 165
Db 121 NKCDVNDKRVQSGEKGLALDYGKFMETSAKANINVENAFTRL 165

RESULT 11
YPT2_SCHPO
ID YPT2_SCHPO STANDARD; PRT; 200 AA.
AC P17609;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein ypt2.
GN YPT2 OR SPAC9B9.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=90269232; PubMed=2112089;
RA Hengst L., Lehmeier T., Gallwitz D.;
RT "Structural and functional analysis of ypt2, an essential ras-related
RT gene in the fission yeast Schizosaccharomyces pombe encoding a Sec4
RT protein homologue.";
RL EMBO J. 9:1957-1962(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90332438; PubMed=2115995;
RA Fawell E., Hook S., Sweet D., Armstrong J.;
RT "Novel YPT1-related genes from Schizosaccharomyces pombe.";
RL Nucleic Acids Res. 18:4264-4264(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

```

```

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Weijtens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- MISCELLANEOUS: THIS PROTEIN IS ESSENTIAL FOR CELL VIABILITY.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC PROBABLE YEAST SEC4 HOMOLOG.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52469; CAA36707.1; -.
DR EMBL; X52864; CAA37045.1; -.
DR EMBL; Z99262; CAB16405.1; -.
DR PIR; S10493; S10493.
DR PIR; S12790; S12790.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_gtp; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (PROBABLE).
FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 200 AA; 22755 MW; 2C658D153A290C30 CRC64;

Query Match 46.2%; Score 510; DB 1; Length 200;
Best Local Similarity 51.9%; Pred. No. 2.6e-34;
Matches 97; Conservative 39; Mismatches 43; Indels 8; Gaps 1;

QY 3 KOYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVRIQ 62
Db 3 KOYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVRIQ 62
QY 4 KSYDVLKLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVRIQ 63
Db 4 KSYDVLKLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVRIQ 63
QY 63 DFAGQERYQTITKQYRRAQIGFLVYDISSERYOHIMKMWSDVDEYAPGVOKILIG 122
Db 63 DFAGQERYQTITKQYRRAQIGFLVYDISSERYOHIMKMWSDVDEYAPGVOKILIG 122
QY 64 DFAGQERYQTITKQYRRAQIGFLVYDISSERYOHIMKMWSDVDEYAPGVOKILIG 123
Db 64 DFAGQERYQTITKQYRRAQIGFLVYDISSERYOHIMKMWSDVDEYAPGVOKILIG 123
QY 123 ADEEQKRVGREGOQOLAKRYGMDVETSACTNLNIKESFTRLTLVLAHRKELEGRLM 182

```



```
DR PIR; JQ0908; JQ0908.
DR PIR; JS0640; JS0640.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfmrng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Multigene family.
FT NP_BIND 22 29 GTP (BY SIMILARITY).
FT NP_BIND 70 74 GTP (BY SIMILARITY).
FT NP_BIND 128 131 GTP (BY SIMILARITY).
FT DOMAIN 44 52 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 213 213 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 214 214 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 216 216 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 216 AA; 23935 MW; 09E9C19A7A44E705 CRC64;

Query Match 46.0%; Score 508; DB 1; Length 216;
Best Local Similarity 46.9%; Pred. No. 4.2e-34;
Matches 100; Conservative 41; Mismatches 54; Indels 18; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDFKMKTEVDGKIVRIQIWDT 64
DB 12 YDLIKLLIGDSGVGKSCLLRFSDGFTTSFTITIGDFKIRTIELDKRIKLIQIWDT 71
QY 65 AGQERYQITTKYYRRAOGIFLVYDISSERSYOHIMKMWSDVDYAVEGVQKILIGNKAD 124
DB 72 ACQERFRITTYAYRGAGIILVYDITDESSFNIRNIRIEQHASDNVNKILVGNKAD 131
QY 125 -BEQRQVREGQOQLAKEYGMDFYETSACTNLNKESTFRTLTELVLQAHKKEGLELRMR 183
DB 132 MDESRAVPTAKGQALADEYGIKFFETSAKTNLNVVEFFSIG-----RDIQR 180
QY 184 ASNELALAE-----LEEEGKPEGPANSSKTC 210
DB 181 LSDTDSRAEPATIKISOTDAQAGQATQKSAC 213

RESULT 14
YPT2_VOLCA
ID YPT2_VOLCA STANDARD; PRT; 217 AA.
AC P36861;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE GTP-binding protein yptv2.
GN YPTV2.
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=f. Nagariensis / HK10;
RX MEDLINE=94037148; PubMed=8221932;
RA Fabry S., Jacobsen A., Huber H., Palme K., Schmitt R.;
RT "Structure, expression, and phylogenetic relationships of a family of
RT ypt genes encoding small G-proteins in the green alga Volvox
RT carteri.";
RL Curr. Genet. 24:229-240(1993).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC -!- TRAFFIC (BY SIMILARITY).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; L08128; AAA34251.1; .
DR PIR; S36365; S36365.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfmrng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport;
FT NP_BIND 20 27 GTP (BY SIMILARITY).
FT NP_BIND 68 72 GTP (BY SIMILARITY).
FT NP_BIND 126 129 GTP (BY SIMILARITY).
FT DOMAIN 42 50 EFFECTOR REGION (PROBABLE).
FT LIPID 215 215 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 216 216 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 217 AA; 24143 MW; 87D3B30C75689EAA CRC64;

Query Match 45.7%; Score 505.5; DB 1; Length 217;
Best Local Similarity 48.6%; Pred. No. 6.7e-34;
Matches 101; Conservative 40; Mismatches 64; Indels 3; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDFKMKTEVDGKIVRIQIWDT 64
DB 10 YDALIKLLVDSGVGKSCLLRFTDMDFTTSFTITIGDFKIRKVDVVDGKILQIWDT 69
QY 65 AGQERYQITTKYYRRAOGIFLVYDISSERSYOHIMKMWSDVDYAVEGVQKILIGNKAD 124
DB 70 ACQERFRITTYAYRGAGIILVYDITDEASFNVRNVRNIRIEQHASDNVNKILVGNKLD 129
QY 125 -BEQRQVREGQOQLAKEYGMDFYETSACTNLNKESTFRTLTELVLQAHKKEGLELRMR 183
DB 130 LAEDKRVYSIARGQALADEFGFRFYETSADKNVHVEEAFIAVAKDVLARMEGEHANQQL 189
QY 184 ASNELALAE-LEEEGKPEGPANSSKTC 210
DB 190 QOQQLSAAQPVRLTSGSP-SPAQGRKSCC 216

RESULT 15
RE10_RAT
ID RB10_RAT STANDARD; PRT; 200 AA.
AC P35281;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; Tissue=Brain;
RX MEDLINE=92210533; PubMed=1313420;
RA Elferink L.A., Anzai K., Scheller R.H.;
RT "Rab15, a novel low molecular weight GTP-binding protein specifically
RT expressed in rat brain.";
RL J. Biol. Chem. 267:5768-5775(1992).
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC -!- NEUROTRANSMITTER RELEASE.
CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN NEURAL AND MUSCLE TISSUES.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; M83677; AAA41991.1; *.
CC PIR; B42148; B42148.
CC HSP; P05713; 3RAB.
CC InterPro; IPR003579; GTPase_Rab.
CC InterPro; IPR001806; Ras_trnsfrmng.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF00071; ras; 1.
CC PRINTS; PR00449; RASTRNSFRMNG.
CC SMART; SM00175; RAB; 1.
CC TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 54 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 200 AA; 22858 MW; EFAP804FD4C3B6A3 CRC64;

Query Match 45.5%; Score 502.5; DB 1; Length 200;
Best Local Similarity 53.5%; Pred. No. 1.1e-33;
Matches 92; Conservative 47; Mismatches 32; Indels 1; Gaps 1;

QY 1 MAKQ-YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVR 59
DB 1 MAKTYDLLFKLLIGDSGVGKTCVLFREFSDAFNTFTSTIEIDFKITVELQGKKIL 60

QY 60 QIWDTAGQERYQTITQYVRRAGIFLVYDISSERYQHIMKWSDVDVEYAPGVOKILI 119
DB 61 QIWDTAGQERFHTITTSYVYGANGIMLVYDITNGKSFENISKWLNRIDQHANEDVERMLL 120

QY 120 GNKADEEQKRVGREGQOQLAKYGMDFYETSACTNLNIKESFTRLTELVLQ 171
DB 121 RNKCDMDHKRVKPKGQEIAREHRIREFETSAKANINIEKAPLTPEDILR 172

Search completed: November 17, 2002, 20:39:12
Job time : 26 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: November 17, 2002, 20:35:44 ; Search time 43 Seconds
(without alignments)
473.965 Million cell updates/sec

Title: US-09-817-198A-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEKGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1077	97.5	212	2 F42148	GTP-binding protei
2	559	50.6	224	2 T33855	hypothetical prote
3	545	49.3	207	2 B49647	GTP-binding protei
4	545	49.3	207	2 B36364	GTP-binding protei
5	532	48.1	203	2 A49647	GTP-binding protei
6	531.5	48.1	200	2 A38625	GTP-binding protei
7	530.5	48.0	200	2 B36364	GTP-binding protei
8	530	48.0	209	2 B38625	GTP-binding protei
9	526	47.6	206	2 I78851	GTP-binding protei
10	522.5	47.3	200	2 B42148	GTP-binding protei
11	519	47.0	216	2 T45901	GTP-binding protei
12	514	46.5	201	2 T28971	hypothetical prote
13	512	46.3	216	2 S33900	GTP-binding protei
14	512	46.3	216	2 T48378	GTP-binding protei
15	510	46.2	200	2 S12790	GTP-binding protei
16	509	46.1	215	2 T14565	GTP-binding protei
17	508.5	46.0	215	2 S7478	GTP-binding protei
18	508	46.0	216	2 J50640	GTP-binding protei
19	507.5	45.9	222	2 T14405	small GTP-binding
20	506	45.8	216	2 S57471	GTP-binding protei
21	505.5	45.7	217	2 S36365	GTP-binding protei
22	502.5	45.5	215	2 S7462	GTP-binding protei
23	501	45.3	203	2 S14955	GTP-binding protei
24	495.5	44.8	204	2 JC7589	Sec4p homolog - ye
25	492.5	44.6	215	2 S57474	GTP-binding protei
26	483.5	43.8	208	2 A34716	GTP-binding protei
27	483	43.7	208	2 A38202	GTP-binding protei
28	482	43.6	203	2 S34253	GTP-binding protei
29	480	43.4	202	2 S38740	GTP-binding protei

30	479	43.3	203	2 B34716	GTP-binding protei
31	478	43.3	202	2 S72515	GTP-binding protei
32	477.5	43.2	201	2 JC2488	GTP-binding protei
33	477.5	43.2	203	2 JC4105	GTP-binding protei
34	477	43.2	201	2 JC5337	GTP-binding protei
35	477	43.2	202	2 S41430	GTP-binding protei
36	476	43.1	258	2 B86153	ARA-5 [imported] -
37	475.5	43.0	203	2 JC1247	GTP-binding protei
38	475	43.0	203	2 S30096	GTP-binding protei
39	473.5	42.9	205	2 T33781	hypothetical protei
40	473	42.8	206	2 T14391	GTP-binding protei
41	473	42.8	210	2 T18242	ras protein homolo
42	470.5	42.6	203	2 B38202	GTP-binding protei
43	470.5	42.6	215	1 TVBYQ4	GTP-binding protei
44	469	42.4	205	2 S38339	GTP-binding protei
45	468	42.4	205	1 TVRTYP	GTP-binding protei

ALIGNMENTS

RESULT 1

F42148

GTP-binding protein rab15 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 02-Feb-2001

C:Accession: F42148

R:Elferink, L.A.; Anzai, K.; Scheller, R.H.

J. Biol. Chem. 267, 5768-5775, 1992

A>Title: rab15, a novel low molecular weight GTP-binding protein specifically express

A:Reference number: A42148; MUID:92210533; PMID:1313420

A:Accession: F42148

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-212 <ELF>

C:Cross-references: GB:M33679; MID:g206536; PIDN:AAA41995.1; PID:g206537

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:210,212/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 97.5%; Score 1077; DB 2; Length 212;

Best Local Similarity 97.6%; Pred. No. 9.2e-79;

Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIYVDGIKVRIQ 60

Db 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMKTIYVDGIKVRIQ 60

QY 61 IWDTAGOERYQITKQYRRAQGIYLVYDISSERSYQHIMKWVDVDEYAPGVQKILIG 120

Db 61 IWDTAGOERYQITKQYRRAQGIYLVYDISSERSYQHIMKWVDVDEYAPGVQKILIG 120

QY 121 NKADEQKRVGREGQQQLAKYGMDFYTSACTNLNINIKESFTRLTETELVLAHRKEGL 180

Db 121 NKADEQKRVGREGQQQLAKYGMDFYTSACTNLNINIKESFTRLTETELVLAHRKEGL 180

QY 181 RMRASNEIALALEEEEGKPEGPANSSKTCWC 212

Db 181 RTCASNELALALEEDEGKTEGPANSSKTCWC 212

RESULT 2

T33855

hypothetical protein D1037.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-Jan-2000

C:Accession: T33855

R:Ledwith, J.; Biewald, T.

submitted to the EMBL Data Library, November 1998

A:Description: The sequence of C. elegans cosmid D1037.

A:Reference number: Z21424
A:Accession: T33855
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-224 <LED>
A:Cross-references: EMBL:AF106592; PIDN:AAC78494.1; GSPDB:GN00019; CESP:D1037.4
A:Experimental source: strain Bristol N2; clone D1037
C:Genetics:
A:Gene: CESP:D1037.4
A:Map position: 1
A:Introns: 10/3; 62/2; 82/3; 181/1
C:Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 50.6%; Score 559; DB 2; Length 224;
Best Local Similarity 55.3%; Pred. No. 2.1e-37;
Matches 105; Conservative 42; Mismatches 39; Indels 4; Gaps 1;
QY 1 MAKQYDLVRLLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMKITVDGKIKVRIQ 60
DB 1 MAKTYDYLKLLLLIGDSGVGKTCVLFREFSDFNNSFISTIGIDFKIRTIELDGKKIKIQ 60
QY 61 IWDTAGQERYQTITKQYRRAGQIFLVYDISSERSYQHIMKWWSDVDEYAPGVQKILIG 120
DB 61 IWDTAGQERFTITTTAYYRGAMGIMLVYDITNERSFENIKNIRNIEEHAASDVVERMIIG 120
QY 121 NKADEQKROVGRGOQOLAKKEYGMDFYETSACTNLNIKESFTRLTFLVLAQHRKELEGL 180
DB 121 NKCDIERREVRSDRGEQLAEVGTFLKTSKANLNIDEAFTFLARDI---KSKMEQN 176
QY 181 RMRASNELAL 190
DB 177 EMRAATGAAI 186

RESULT 3

B49647
GTP-binding protein rab8 - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Jan-2001
C:Accession: B49647; S36817
R:Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavitian, A.; Louvar
J. Cell Biol. 124, 101-115, 1994
A:Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cell
A:Reference number: A49647; MUID:94124602; PMID:8294494
A:Accession: B49647
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-207 <ZAH>
A:Cross-references: EMBL:X56741; NID:g452317; PIDN:CAA40065.1; PID:g452318
R:Joberty, G.; Tavitian, A.; Zahraoui, A.
FEBS Lett. 330, 323-328, 1993
A:Title: Isoprenylation of Rab proteins possessing a C-terminal Caax motif.
A:Reference number: S36817; MUID:93387463; PMID:8375503
A:Accession: S36817
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 175-186 <JOB>
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: Gtp binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;
F:9-124/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:121-124/Region: GTP-binding NKXD motif
F:151-153/Region: GTP-binding SAK/L motif
F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 49.3%; Score 545; DB 2; Length 207;
Best Local Similarity 52.9%; Pred. No. 2.4e-36;
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;
QY 1 MAKQYDLVRLLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMKITVDGKIKVRIQ 60
DB 1 MAKTYDYLKLLLLIGDSGVGKTCVLFREFSDFNNSFISTIGIDFKIRTIELDGKKIKIQ 60

QY 61 IWDTAGQERYQTITKQYRRAGQIFLVYDISSERSYQHIMKWWSDVDEYAPGVQKILIG 120
DB 61 IWDTAGQERFTITTTAYYRGAMGIMLVYDITNERSFENIKNIRNIEEHAASDVVERMIIG 120
QY 121 NKADEQKROVGRGOQOLAKKEYGMDFYETSACTNLNIKESFTRLTFLVLAQHRKELEGL 180
DB 121 NKCDVNDKROVSKERGEKALDYGKFMETSAKANINVENAFTFLARDIKAKMDKKLEGN 180
QY 181 RMRASNE 187
DB 181 SPQGSNQ 187
RESULT 4
B36364
GTP-binding protein rab8 - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 02-Feb-2001
C:Accession: B36364; S15604
R:Chavrier, P.; Vangron, M.; Sander, C.; Simons, K.; Zerial, M.
Mol. Cell. Biol. 10, 6578-6585, 1990
A:Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.
A:Reference number: A36364; MUID:91061765; PMID:2123294
A:Accession: B36364
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-207 <CHA>
A:Cross-references: GB:X56385; NID:9920; PIDN:CAB56776.1; PID:g6006436
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: Gtp binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop
F:9-124/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:121-124/Region: GTP-binding NKXD motif
F:151-153/Region: GTP-binding SAK/L motif
F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 49.3%; Score 545; DB 2; Length 207;
Best Local Similarity 52.9%; Pred. No. 2.4e-36;
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDLVRLLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMKITVDGKIKVRIQ 60
DB 1 MAKTYDYLKLLLLIGDSGVGKTCVLFREFSDFNNSFISTIGIDFKIRTIELDGKKIKIQ 60
QY 61 IWDTAGQERYQTITKQYRRAGQIFLVYDISSERSYQHIMKWWSDVDEYAPGVQKILIG 120
DB 61 IWDTAGQERFTITTTAYYRGAMGIMLVYDITNERSFENIKNIRNIEEHAASDVVERMIIG 120
QY 121 NKADEQKROVGRGOQOLAKKEYGMDFYETSACTNLNIKESFTRLTFLVLAQHRKELEGL 180
DB 121 NKCDVNDKROVSKERGEKALDYGKFMETSAKANINVENAFTFLARDIKAKMDKKLEGN 180
QY 181 RMRASNE 187
DB 181 SPQGSNQ 187
RESULT 5
A49647
GTP-binding protein Rab13 - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 02-Feb-2001
C:Accession: A49647
R:Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavitian, A.; Lou
J. Cell Biol. 124, 101-115, 1994
A:Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized c
A:Reference number: A49647; MUID:94124602; PMID:8294494
A:Accession: A49647
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-203 <ZAH>
A:Cross-references: EMBL:X75593; NID:g452319; PIDN:CAA53266.1; PID:g452320
C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C;Species: Canis lupus familiaris (dog)
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 02-Feb-2000

D6 121 NRCDMNEXRQVSKERGEKLAIDYGKF-ETSAKSSINVEEAF-ITLARDIMTKLNKMM----

QY	180	LRMRASNELALAELEEEEGKPEGPANSSKTCWC	212
		: : :	
Db	172	AKPFDSTD---EQSRDTVPNPVQPRQSSSGGC	200

Db	12	YDYLIKULLIGDGVGSKLLRSDDFTTFTTIGIDFKIRVELDGKKIKIQT	71
Qy	65	AGQRYOTITQYYRRAQIGFLVYDISSEYOHIMKWSDVDEYAPGVQKILGNKAD	124
Db	72	AGQRFRTITAYYRGAMGILLVVDVDTDESFNIRNMKNIEQASDNVKILVGNKAD	131
Qy	125	-EEOKRVGREGQOOLAKHEYGMDFEYETSACTNLNIKESF-----TRLTELVLQAHRK	175
Db	132	MDESKRAVPTAKQALADEYGIKFETSAKTNLVNENVFMSIAKDIKQRLTETDIKA---	188
Qy	176	ELEGLMRASNELALAELEE	196
Db	189	EPOGIKTTKODPTAASSSTAER	209

DB 189 EPQGIKRQKQDTAAASSSTAER 209

RESULT 15

S12790

GTP-binding protein ypt2 - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 02-Feb-2001

C:Accession: S12790; S10493; T39214

R:Haubruck, H.; Engelke, U.; Mertins, P.; Gallwitz, D.

EMBO J. 9, 1957-1962, 1990

A:Title: Structural and functional analysis of ypt2, an essential ras-related gene i

A:Reference number: S12790; PMID:90269232; PMID:2112089

EMBL U0197, 1997, 43042; ID: 197
A>Title: Structural and functional analysis of ypt2, an essential ras-related gene 1
A:Reference number: S12790; MUID:90269232; PMID:2112089
A:Accession: S12790
A:Molecule type: DNA
A:Residues: 1-200 <HAU>
A:Cross-references: EMBL:X52469; NID:g5143; PIDN:CRAA36707.1; PID:g5144
R:Fawell, E.; Hook, S.; Sweet, D.; Armstrong, J.
Nucleic Acids Res. 18, 4264, 1990
A>Title: Novel YPT1-related genes from Schizosaccharomyces pombe.
A:Reference number: S10492; MUID:90332438; PMID:2115995
A:Accession: S10493
A:Molecule type: DNA
A:Residues: 1-200 <FAW>
A:Cross-references: EMBL:X52864; NID:g5149; PIDN:CRAA37045.1; PID:g5150
R:McDougall, R.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21836
A:Accession: T39214
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2002, 20:42:04 ; Search time 51 seconds
(without alignments)
1536.011 Million cell updates/sec

Title: US-09-817-198A-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEKGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328367 seqs, 184756068 residues

Total number of hits satisfying chosen parameters: 656734

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09817198/runat_13112002_135207_7068/app_query.fasta_1.391
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09817198@cgn.1.1.20 @runat_13112002_135207_7068
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LOGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1105	100.0	3257	10	US-09-817-198A-1
2	534.5	48.4	1274	10	US-09-925-302-91
3	530.5	48.0	1537	10	US-09-925-300-631
4	528.5	47.8	624	10	US-09-794-257-9
					Sequence 1, Appli
					Sequence 91, Appl
					Sequence 631, App
					Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-09-817-198A-1

; Sequence 1, Application US/09817198A

; Patent No. US20020146758A1

; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; FILE REFERENCE: CLO01188

; CURRENT APPLICATION NUMBER: US/09/817.198A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: fastseq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 3257

; TYPE: DNA

; ORGANISM: Human

US-09-817-198A-1

Alignment Scores:

Pred. No.: 2.14e-129 Length: 3257
Score: 1105.00 Matches: 212
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

Sequence 7, Appli
Sequence 879, App
Sequence 885, App
Sequence 894, App
Sequence 896, App
Sequence 836, App
Sequence 4, Appli
Sequence 529, App
Sequence 832, App
Sequence 1461, Ap
Sequence 10, Appl
Sequence 13, Appl
Sequence 1, Appli
Sequence 66, Appl
Sequence 15, Appl
Sequence 49, Appl
Sequence 794, App
Sequence 2113, Ap
Sequence 478, App
Sequence 49, Appl
Sequence 340, App
Sequence 646, App
Sequence 1359, Ap
Sequence 1, Appli
Sequence 78, Appl
Sequence 652, App
Sequence 563, App
Sequence 374, App
Sequence 774, App
Sequence 469, App
Sequence 1231, Ap
Sequence 1426, Ap
Sequence 4, Appli
Sequence 6, Appli
Sequence 1703, Ap
Sequence 1703, Ap
Sequence 4, Appli
Sequence 235, App
Sequence 901, App
Sequence 3393, Ap
Sequence 1, Appli

[illegible]

Mon Nov 18 09:05:56 2002

```

; PRIOR FILING DATE: 2000-04-14
;
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 896
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-896

```

Alignment Scores:		
Pred. No.:	6.62e-57	2497
Score:	528.50	95
Percent Similarity:	80.68%	Conservative: 47
Best Local Similarity:	53.98%	Mismatches: 33
Query Match:	47.83%	Indels: 1
DB:	10	Gaps: 1

US-09-817-198A-2 (1-212) x US-09-834-975-896 (1-2497)

Qy	1	MetAlaLysGlnTyrAspValLeupheArgLeuLeuLeuIleGlyAspSerGlyValGly	20
Db	92	ATGGCGAAGACGTACGATTATCTCTTCAAGCTCCTGCTGATGGGACTCGGGGTAGGC	151
Qy	21	LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr	40
Db	152	AAGACCTTCCTCTGTTCCCTTCTCAGAGGACGCCCTTCAACACCACCTTCATCTCCACC	211
Qy	41	IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValargIleGln	60
Db	212	ATCGCAATTGATTTAAATATAGACGATAGAACTAGATGGGAAGAAAATTAAGCTTCAG	271
Qy	61	IleTrpAspThrAlaGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAsp	80
Db	272	ATATGGGACACAGGGGTACGAAAGATTCGGAAACAACTACGACACCGCTACTACAGAGA	331
Qy	81	AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet	100
Db	332	GCCATGGGCATTATGCTGGTCTATGACATCACAAATGAAAATCCCTTGACAAATATAAA	391
Qy	101	LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly	120
Db	392	AATTGGATCAGAACATTGAAGACATGCCTCTTCCGATGTCGAAGAAGATGATCCGGGT	451
Qy	121	AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla	140
Db	452	NACAAATGTGATATGATGACAAAAGACAAGTGTCAAAAGAAAAGAGGGGAGAGAGCTAGCA	511
Qy	141	LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu	160
Db	512	ATTGACTATGGGATTAATCTTCGGAGCAAGCGCAAAATCCCAATTTGATAGAGAG	571
Qy	161	Ser---PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLys	175
Db	572	GCATTTTTTACCTTCGACGAGATATATGACAAACCTCAACAGAAAA	619

RESULT 10

US-09-938-842A-836
; Sequence 836, Application US/09938842A
; Patent No. US20020160378A1

; FACILE NO. US2002010
: GENERAL INFORMATION:

: APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

; TITLE OF INVENTION: S

; TITLE OF INVENTION:

; FILE REFERENCE: SCRIPT:

[illegible]

; CURRENT FILING DATE:

2

```

: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPT300-3
: CURRENT APPLICATION NUMBER: US/09/938,842A
: CURRENT FILING DATE: 2001-08-24

```

```

, PRIOR APPLICATION NUMBER: US 60/227,866
, PRIOR FILING DATE: 2000-08-24
, PRIOR APPLICATION NUMBER: US 60/264,647
, PRIOR FILING DATE: 2001-01-16
, PRIOR APPLICATION NUMBER: US 60/300,111
, PRIOR FILING DATE: 2001-06-22
, NUMBER OF SEQ ID NOS: 5379
, SEQ ID NO 836
, LENGTH: 651
, TYPE: DNA
, ORGANISM: Arabidopsis thaliana
US-09-938-842A-836

```

Alignment Scores:	1.51e-56	Length:	651
Pred. No.:	539.00	Matches:	102
Score:	7.64%	Conservative:	42
Percent Similarity:	50.75%	Mismatches:	42
Best Local Similarity:	50.75%	Indels:	15
Query Match:	46.97%	Gaps:	3
DB:	9		

US-09-817-198A-2 (1-212) x US-09-938-842A-836 (1-651)

[illegible]

RESULT 11

US-09-967-736-4

; Sequence 4, Application US/09967736

Patent No. US20020103340A1

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

Lal, Preeti
Corley, Neil C.
Shah, Purvi
TITLE OF INVENTION: RAB PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICANT: FastSeq for Windows Version 2.0
FILING DATE: 28-Sep-2001
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/154,602
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0367 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-853-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 925 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LIVRUT04
CLONE: 2514506
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-967-736-4

Alignment Scores:
Pred. No.: 3,18e-51 Length: 925
Score: 478.50 Matches: 91
Percent Similarity: 65.84% Conservative: 42
Best Local Similarity: 45.05% Mismatches: 64
Query Match: 43.30% Indels: 5
DB: 10 Gaps: 1

US-09-817-198a-2 (1-212) x US-09-967-736-4 (1-925)

QY 1 MetaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly 20
Db 66 ATGAACCCGAAATGACTACCTGTTTAAAGCTGCTTTGATTGGGAGCTCAGCGGTGGGC 125
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisLeuSerThr 40
Db 126 AAGTCATGCTGCTCTCTGGGTTTGGCTGATGACAGCTACACAGAGCTACATCAGCACC 185
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 186 ATCGGGTGGACTTCAGATCCGACCATCGAGCTGGGAGGAACTATCAAACTTCAG 245
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleLysGlnTyrTyrArgArg 80
Db 246 ATCTGGGACACAGCGGGCCAGGAAGGTTCCGGACCATCACTTCCAGCTACTACCGGGG 305
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 306 GCTATGGCATCATCTGCTGTATACGCTACGACAGGAAATCTACGCCAACCTGAAG 365
QY 101 LysTrpValSerAspValAspGlnTyrAlaProGluGlyValGlnLysIleLeuGly 120

Db 366 CAGTGGCTGCAGGAGATTGACCGCTATGCCAGCGAGAGCTCAATAAGCTCTGGTGGGC 425
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGlnGlnGlyClnGlnLeuAla 140
Db 426 AACAAAGAGCGACCTCACCACCAAGAGGTGGGACAAACACACACCAAGAGTTTGA 485
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 486 GACTCTCTGGGATCCCTCTTGGAGAGAGCGGCAAGAAATGCCACCAATGTCGAGCAG 545
QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGlnLeuGlyLeu 180
Db 546 GCGTTC-----ATGACCATGCTGCTGAAATCAAAAGCGGATGGGCGCT 590
QY 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysPro 200
Db 591 GGAGCAGCCTCTGGGGGCGAGCGGCCCAATCTCAAGATGCGACAGCACCCCTGTAAGCG 650
QY 201 GluGly 202
Db 651 GCTGGC 656

RESULT 12

US-09-770-445-529
Sequence 529, Application US/09770445

Patent No. US20020023281A1
GENERAL INFORMATION:

APPLICANT: Gorlach, Jorn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Matthew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurdan, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2023US (PARA-012PRV)
CURRENT APPLICATION NUMBER: US/09/770,445
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: US 60/178,472
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 529
LENGTH: 881
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-445-529

Alignment Scores:
Pred. No.: 3,96e-51 Length: 881
Score: 477.50 Matches: 96
Percent Similarity: 65.57% Conservative: 43
Best Local Similarity: 45.28% Mismatches: 55
Query Match: 43.21% Indels: 18
DB: 10 Gaps: 4

US-09-817-198a-2 (1-212) x US-09-770-445-529 (1-881)

QY 1 MetaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly 20

Db	60	ATGAATCCTCAGTACGACTATCTTTTCAAGCTCCGCTTATCGGGATCTCGGATAGGC	119
QY	21	LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr	40
Db	120	AAGTCGTGCTCTTTTGAGATCTCTGATGATCTTATGTAGAAAGTTACATTAGCACT	179
QY	41	IleGlyValAspPheIysMetLysThrIleGluValAspGlyIleIysValargIleGln	60
Db	180	ATTGGAGTCGATTTAAAAATTAGGACTGTGGACAAAGATTGCAAAACAATTAAGCTCAA	239
QY	61	IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg	80
Db	240	ATTGGGACACTGCTGCTCAAGAACGGTTCAGGACTATTACTAGCAGTTACTACCTGGG	299
QY	81	AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet	100
Db	300	GCACATCGGAATTATTATTGCTACGATGTCACAGATCAAGAAGCTCAATAATGCTCAAG	359
QY	101	LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly	120
Db	360	CAATGGTTGAGTGAATGATCGTTATGCTAGTCAGCAANTGTCACAAACTCCTTGTTGA	419
QY	121	AsnLysAlaAspGluGlnLysArgGlnValIcIyargGluGlnGlyGlnGlnLeuAla	140
Db	420	AACAAGCTGATCTTACTGAAACACAGAGCACCATTCCCTATGAAACTGCCAAGCGTTTGGC	479
QY	141	LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu	160
Db	480	GATGAATCGGGATTCCTTTATGGAGACTAGTGTCAAAAGATGCTACAACGTAGAACAG	539
QY	161	SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu	180
Db	540	GCTTTC-----ATGGCAATCTCTGCATCCATCAAGAG-----	572
QY	181	ArgMetArgAlaSer-AsnGluLeuAlaLeuAlaGluLeuGluGluGluGluGlyLysPr	200
Db	573	AGAATG---GCTAGCCAAACAGCTGGGAATAATGCAA-----GA	608
QY	200	oGluGlyProAlaAsnSerSerLysThrCysTrp	211
Db	609	CCACGCGCTGTCAGATCAGAGGACAGCCCTGTGG	642

RESULT 13
US-09-938-842A-832
; Sequence 832, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRI1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 832
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-832

Alignment Scores:	
Pred. No.:	3,67e-50
Score:	468.00
Length:	609
Matches:	84

2.

[illegible]

```

1 RESULT 14
2 US-09-917-800A-1461
3 : Sequence 1461, Application US/09917800A
4 : Patent No. US20020119462A1
5 :
6 : GENERAL INFORMATION:
7 : APPLICANT: Mendrick, Donna
8 : APPLICANT: Porter, Mark
9 : APPLICANT: Johnson, Kory
10 : APPLICANT: Castle, Arthur
11 : APPLICANT: Elashoff, Michael
12 : APPLICANT: Gene Logic, Inc.
13 :
14 : TITLE OF INVENTION: Molecular Toxicology Modeling
15 :
16 : FILE REFERENCE: 44921-5038-US
17 : CURRENT APPLICATION NUMBER: US/09/917,800A
18 : CURRENT FILING DATE: 2001-07-31
19 :
20 : PRIOR APPLICATION NUMBER: US 60/222,040
21 : PRIOR FILING DATE: 2000-07-31
22 :
23 : PRIOR APPLICATION NUMBER: US 60/222,880
24 : PRIOR FILING DATE: 2000-11-02
25 :
26 : PRIOR APPLICATION NUMBER: US 60/290,029
27 : PRIOR FILING DATE: 2001-05-11
28 :
29 : PRIOR APPLICATION NUMBER: US 60/290,645
30 : PRIOR FILING DATE: 2001-05-15
31 :
32 : PRIOR APPLICATION NUMBER: US 60/292,336
33 : PRIOR FILING DATE: 2001-05-22
34 :
35 : PRIOR APPLICATION NUMBER: US 60/295,798
36 : PRIOR FILING DATE: 2001-06-06
37 :
38 : PRIOR APPLICATION NUMBER: US 60/297,457
39 : PRIOR FILING DATE: 2001-06-13
40 :
41 : PRIOR APPLICATION NUMBER: US 60/298,884

```

```

/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: US
/ ZIP: 94304
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/051,986
/ FILING DATE: 15-Jan-2002
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/766,551
/ FILING DATE: DECEMBER 12, 1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Cerrione, Michael C.
/ REGISTRATION NUMBER: 39,132
/ REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-855-0555
/ TELEFAX: 650-845-4166
/ TELEX: <Unknown>
/
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 875 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: UCMCL5T01
/ CLONE: 1528559
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 10 :
US-10-051-986-10

Alignment Scores:
Pred. No.: 5,086-45 Length: 875
Score: 429.00 Matches: 82
Percent Similarity: 68.48% Conservative: 44
Best Local Similarity: 44.57% Mismatches: 46
Query Match: 38.82% Indels: 12
DB: 12 Gaps: 3

US-09-817-198A-2 (1-212) x US-10-051-986-10 (1-875)
QY 5 TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24
Db 72 TAGCACTTCACGGGCAAGGTGATGCTCTCGGAGACACAGGGCTCGGCAAAACATGTTTC 131
QY 25 LeuCysArgPheThrAspAsnGluPheHisSer--SerHisIleSerThrIleGlyVal 43
Db 132 CTGATCCAAATTCAAAGACGGGGCTTCCTGTCGGGAACCTTCATAGCCACCGTCGGSCATA 191
QY 44 AspPheLysMetLysThrIleGluValaspGlyIleLysValArgIleGlnIletrpasp 63
Db 192 GACTTCAGAAACAAGGTGGTGACTGTGTGGATGCGGTGAGAGTGAAGCTGCAGATCTGGGAC 251
QY 64 ThrAlaGlyClnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGly 83
Db 252 ACCGCTGGCGAAGAACGGTTCGAAACGGTCACCCATGCTATTACAGAGATGCTCAGGCC 311
QY 84 IlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpVal 103
Db 312 TTGCTTCTGCTGTATGACATCATCCACAACAAATCTTCTTCGCAACACATCAGGCCCTGGCTC 371
QY 104 SerAspValaspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAla 123
Db 372 ACTGACATTCATGAGTATGCCAGAGGGGACGTGGTGATCATCTGCTAGGCAACAGGGC 431

```

```

QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnGlnLeuAlaLysGluTyr 143
    ||| :||| :||| ||| |||:| |||:| |||:| |||:| |||:|
Db 432 GATATGAGCAGCGAAGAGTGATCCGTTCCGAAGACGGAGAGACCTTGCCAGGGAGTAC 491
    |||:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163
    |||:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 GGTGTTCCCTTCCTGGAGACCAGCGCAAGACTGGCATGAATGTG----- 536
    |||:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 -----GAGTTAGCCTTTCTGGCCATCGCCACAGGAA-----CTGAATACCAG 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 AlaSerAsnGlu 187
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 579 GCCGGGCATCAG 590
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

Search completed: November 17, 2002, 22:06:14
 Job time : 55 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2002, 20:38:44 ; Search time 255 Seconds
(without alignments)
1872.250 Million cell updates/sec

Title: US-09-817-198a-2

Perfect score: 1105
Sequence: 1 MAKQYDLFRLLIGSGVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09817198/runat_13112002_135206_7022/app_query.fasta_1.391
-DB=N.Geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198 -CGN_1_1_125 -runat_13112002_135206_7022 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1105	100.0	2021	22	CDNA encoding nove
2	1092	98.8	1666	21	Human ORFX ORF1368
3	832	75.3	566	22	CDNA encoding nove
4	546.5	49.5	1540	23	Drosophila melanog
5	545	49.3	1986	22	Human rab8 homolog
6	540	48.9	2411	24	Human ovarian anti
7	534.5	48.4	1274	21	Lung cancer associ
8	530.5	48.0	716	21	Human Rab10 CDNA
9	530.5	48.0	861	21	Human Rab10 CDNA #
10	530.5	48.0	888	21	Canine Rab10 CDNA
11	530.5	48.0	956	21	Nucleotide sequenc
12	530.5	48.0	1537	21	Human prostate can
13	530.5	48.0	3533	22	Human cDNA sequenc
14	529.5	47.9	674	23	DNA encoding novel
15	528.5	47.8	1161	22	Nucleotide sequenc
16	528.5	47.8	2247	23	Drosophila melanog
17	528.5	47.8	2497	22	Human cancer agent
18	528.5	47.8	2497	22	Human cancer agent
19	528.5	47.8	2497	22	Human cancer agent
20	528.5	47.8	2497	22	Human cancer agent
21	528.5	47.8	2497	23	Human prostate exp
22	528.5	47.8	2497	23	Human prostate exp
23	528.5	47.8	3077	22	Human cDNA sequenc
24	519.5	47.0	866	22	Human cDNA clone (
25	519	47.0	1023	21	Arabidopsis thalia
26	519	47.0	1025	21	Arabidopsis thalia
27	513.5	46.5	1129	21	Zea mays DNA fragm
28	510	46.2	911	21	Arabidopsis thalia
29	507	45.9	1101	21	Arabidopsis thalia
30	504	45.6	1203	21	Zea mays DNA fragm
31	501	45.3	705	21	Arabidopsis thalia
32	494	44.7	1153	21	Arabidopsis thalia
33	478.5	43.3	925	20	Human Rab protein,
34	478.5	43.3	939	21	Human breast and o
35	478.5	43.3	1898	22	CDNA encoding nove
36	478.5	43.3	1944	22	Human polynucleoti
37	477.5	43.2	730	21	Nucleotide sequenc
38	477.5	43.2	777	21	Arabidopsis thalia
39	477.5	43.2	881	24	Arabidopsis thalia
40	472.5	42.8	607	22	Peppermint plant o
41	472.5	42.8	1202	21	Aspergillus oryzae
42	468	42.4	959	21	Arabidopsis thalia
43	467	42.3	723	24	Human cDNA differe
44	467	42.3	2528	21	Human secreted pro
45	466	42.2	932	21	Arabidopsis thalia

ALIGNMENTS

RESULT 1

ID AAS27053 standard; cDNA; 2021 BP.

XX AAS27053;

DT 07-NOV-2001 (first entry)

XX CDNA encoding novel signal transduction pathway protein, Seq ID 88.
DE Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;

KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.

XX Homo sapiens.

XX WO200154733-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01312.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 18-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226868.

XX 22-AUG-2000; 2000US-0227182.

XX 23-AUG-2000; 2000US-0227009.

XX 30-AUG-2000; 2000US-0228924.

XX 01-SEP-2000; 2000US-0229287.

XX 01-SEP-2000; 2000US-0229343.

XX 01-SEP-2000; 2000US-0229344.

XX 01-SEP-2000; 2000US-0229345.

XX 05-SEP-2000; 2000US-0229509.

XX 05-SEP-2000; 2000US-0229513.

XX 06-SEP-2000; 2000US-0230437.

XX 06-SEP-2000; 2000US-0230438.

XX 08-SEP-2000; 2000US-0231242.

XX 08-SEP-2000; 2000US-0231243.

XX 08-SEP-2000; 2000US-0231244.

XX 08-SEP-2000; 2000US-0231413.

XX 08-SEP-2000; 2000US-0231414.

XX 08-SEP-2000; 2000US-0232080.

XX 12-SEP-2000; 2000US-0232081.

XX 12-SEP-2000; 2000US-0231968.

XX 14-SEP-2000; 2000US-0232397.

XX 14-SEP-2000; 2000US-0232398.

XX 14-SEP-2000; 2000US-0232399.

XX 14-SEP-2000; 2000US-0232400.

XX 14-SEP-2000; 2000US-0232401.

XX 14-SEP-2000; 2000US-0233063.

XX 14-SEP-2000; 2000US-0233064.

XX 14-SEP-2000; 2000US-0233065.

XX 21-SEP-2000; 2000US-0234223.

XX 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 03-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-463460/50.
 DR P-PSDB; AAU17136.
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders -
 XX Claim 1; SEQ ID No 88; 880pp; English.
 PS The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
 CC pathway protein coding sequences and PCR primers of the invention.
 XX

Alignment Scores:
 Pred. No.: 6.03e-118 Length: 2021
 Score: 1105.00 Matches: 212
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-817-198A-2 (1-212) x AAS27053 (1-2021)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuLeuValGly 20
 Db 81 ATGGCGAGCAGTACGATGTCTGTTCCGGCTGCTGTGATCGGGAGCTCCGGGTGGGC 140
 Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisSerThr 40
 Db 141 AAGACCTGCTGCTGCGGCTTACCGACAACGAGTTCCTCTCGCACATCCACC 200
 Qy 41 IleGlyValAspPheLysMetLysThrIleGluValaspGlyIleLysValArgIleGln 60
 Db 201 ATCGGTGTGATTTAAGATGAAGACCATAGAGGTAGCGGCATCAAGTCCGATACAG 260
 Qy 61 IleTTPAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
 Db 261 ATCTGGGACATGTCAGGGCAGGAGATACACAGACCATCAACAGTACTATCGGGG 320
 Qy 81 AlaGlnGlyIlePheLeuValTyrAspLysSerSerGluArgSerTyrGlnHisIleMet 100
 Db 321 GCCAGGGGATATTTTGGTCTATGACATATGACGAGCGCTCTTACCAGCACATCATG 380
 Qy 101 LysTrpValSerAspValaspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
 Db 381 AAGTGGGTGAGTGTGATGAGTACGCACCAAGGCGTCCAGGAAGATCCTTATTGGG 440

Qy 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
 Db 441 AATAAGGCTGATGAGGAGCAGAAACGGCAGGTGGGAAGAGACCAAGGGCAGCTGGCG 500
 Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 Db 501 AAGGAGTATGGCATGCACTTCTATGAACACAGTGGCTGCACCAACCTCAACATTAAAGAG 560
 Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGlyLeu 180
 Db 561 TCATTACGCGTCTCACAGAGCTGGTGTGTCAGGCCCATAGGAAGAGCTGGGAAGCCCTC 620
 Qy 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysPro 200
 Db 621 CGGATCGCTGCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGAGGAG 680
 Qy 201 GluGlyProAlaAsnSerSerLysThrCysTrpCys 212
 Db 681 GAGGCGCCAGCGCACTTTCGAAAAACCTGCTGGTGC 716
 RESULT 2
 AAC75813
 ID AAC75813 standard; cDNA; 1666 BP.
 XX AAC75813;
 XX 08-FEB-2001 (first entry)
 XX Human ORFX ORF1368 polynucleotide sequence SEQ ID NO:2735.
 DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antiporiatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX Homo sapiens.
 OS
 XX WO200058473-A2.
 XX 05-OCT-2000.
 XX 31-MAR-2000; 2000WO-US08621.
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shimkets RA, Leach M;
 PI
 DR WPI; 2000-602362/57.
 DR P-PSDB; AAB41604.
 XX Novel nucleic acids and peptides derived from open reading frame x,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX Claim 5; Page 1978-1979; 5507pp; English.
 PS AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX

PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234977.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249284.
 PR 17-NOV-2000; 2000US-0249285.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX
 DR WPI; 2001-465460/50.
 DR P-PSDB; AAU17555.
 XX
 PT Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders
 XX
 PS Claim 1; SEQ ID No 507; 880pp; English.
 XX
 CC The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
 CC pathway protein coding sequences and PCR primers of the invention.
 XX

Alignment Scores:
 Pred. No.: 4.6e-87 Length: 566
 Score: 832.00 Matches: 160
 Percent Similarity: 97.56% Conservative: 0
 Best Local Similarity: 97.56% Mismatches: 4
 Query Match: 75.29% Indels: 0
 DB: 22 Gaps: 0

US-09-817-198A-2 (1-212) x AAS27472 (1-566)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
 |||||
 Db 71 ATGGCGAAGCAGTAGCTGTGCTCGGGTGTGCTGTGCTGCGGGGACTCCGGGGTGGGC 130

QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
DB 131 AAGACCTGCGTGTGCGGCTTCCCGACACAGAGTTTCCACTCTCGCACATCTCCACC 190
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
DB 191 ATCGGTGTGACTTTAAGATGAAGACCATAGAGGTAGAGCGGCATCAAGTGGGATACAG 250
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
DB 251 ATCTGGGACACTGCAGGCGAGAGATACACAGACCATCACAAGCAGTACTATCGGGG 310
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
DB 311 GCCCAGGGGATATTTTGGTCTATGATCATAGCAGGCGGCTCTTACCAGCACATCATG 370
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
DB 371 AAGTGGTCACTGACGTGGATGAGTACGCCACCCAGAGGCGTCCAGAAATCCTTATTGGG 430
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140
DB 431 AATAAGGCTGATGAGGAGCAGAWACGCGCAGGTGGGAAGAGAGCAAGGCGCAGCANCCTGGCG 490
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
DB 491 AANGAGTATGGCATGGACTTCTATGAACAAGTGCCTGCACCAACCTNACATTAAAGAG 550
QY 161 SerPheThrArg 164
DB 551 TCATTCACGCGT 562
RESULT 4
ABL27707
ID ABL27707 standard; DNA; 1540 BP.
XX AC ABL27707;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 34594.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Claim 1; SEQ ID NO 34594; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1540 BP; 437 A; 353 C; 420 G; 330 T; 0 other;
Alignment Scores: Length: 1540
Pred. NO.: 2.19e-53 Matches: 105
Score: 546.50 Conservativity: 47
Percent Similarity: 72.04% Mismatches: 50
Best Local Similarity: 49.76% Indels: 9
Query Match: 49.46% Gaps: 23
DB: 23
US-09-817-198A-2 (1-212) x ABL27707 (1-1540)
QY 1 MetAlaLysGln--TyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuLeuLeuVal 19
DB 257 ATGGCAAGAAACCTACGATTGCTTTAAACTGTTCATCGGTGATTTCAGAGTG 316
QY 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
DB 317 GCGAAGACGTGATATTTGTCGGGTTCGGATGATGATTCACGTCACGTTCTATATCG 376
QY 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIle 59
DB 377 ACCATAGGCATCGATTTCAAAATCAAAACAGTCGAGTCGCGGCAAGAGATCAAGCTG 436
QY 60 GlnIleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 79
DB 437 CAAATATGGACACCGCGCGCAGGAGCGGTTCACACAGCATAAACACCTCGTACTATCGA 496
QY 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
DB 497 GCGCCATGGCATATGCTGTCTATGACATAACGAGAGAGAGAGAGAGAGAGAGAGATT 556
QY 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
DB 557 GTCAATGTTACGGAATATAGACGACGCGCAACGAGGATGTGGAGAAGATGATCCTC 616
QY 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu 139
DB 617 GGCACAAGTGCATATGACGGACAAGAGGCGGTTCACAACGAGGAGCGCGCAAGCGATT 676
QY 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
DB 677 GCCCGTGACATGGCATTCGGTTTATGGAACATCCGCCAAGTCGAACATAACATCGAG 736
QY 160 GluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGly 179
DB 737 CGGCGCTTCTCGAGCTGGCGGAGCCATCTGGAC-----AAGACATCAGGA 784
QY 180 LeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGlyLys 199
DB 785 CGCGAGTCGCGGAGAAATCAGGAGCGGTGATTATTCATCGCCGGAACCGAGGAGCGG 844
QY 200 ProGluGlyProAlaAsnSerSerLysThrCys 210
DB 845 CCG-----GGCTACAGCAAGTGTGTC 865
RESULT 5
ABA09160
ID ABA09160 standard; cDNA; 1986 BP.
XX AC ABA09160;
XX DT 11-JAN-2002 (first entry)
XX PR Human rab8 homologue-encoding cDNA, SEQ ID NO:936.
XX DE

PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -
XX
XX Claim 1; Page 566; 1425pp; English.
XX
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.
XX
XX Sequence 1274 BP; 416 A; 267 C; 353 G; 235 T; 3 other;
SQ

Alignment Scores:
Pred. No.: 4.15e-52 Length: 1274
Score: 534.50 Matches: 102
Percent Similarity: 66.38% Conservative: 54
Best Local Similarity: 43.40% Mismatches: 55
Query Match: 48.37% Indels: 25
DB: 21 Gaps: 1

US-09-817-198a-2 (1-212) x AAF18072 (1-1274)

QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
DB 132 ATGGCCAAAGCCATACGACACCTCTTCAAGTTGCTGCTGATCGGGGACTCGGGGTGGC 191

QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisSerThr 40
DB 192 AGACTGTGCTGATCATTCGCTTTCGAGAGGCACTTCAACAACACTTACATCTCCACC 251

QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValGlnGln 60
DB 252 ATCGGAATTGATTCAAGTTCGCACTGTGGATATAGAGGGAAGAGATCAACTACAA 311

QY 61 IleTyrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 80
DB 312 GTCTGGGACACGGCTGGCCCAAGCGGTTCAGACAATAACTACTGCTTACTACCGTGA 371

QY 81 AlaGlnGlyPheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
DB 372 GCATCGGGCATTCCTAGTATACGACATPCACGGATGAGAAATCTTTCAGAAATATTCAG 431

QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
DB 432 AACTGGATGAAGACATCAAGAGAAATGCTCGGCTGGGGTGGAGCGCTCTTCGTGGG 491

QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
DB 492 AACAAATGTGACATGGAGCCCAAGAGAGAGGTGCAGAGGACGAGCCCGATGTTGGCT 551

QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
DB 552 CGAGAGCATGGAATCCGATTTTCGAACTAGTCTGCTAAATCCAGTATGATGATGATGAG 611

QY 161 SerPheThrArgLeuThrGluLeuValLeu----- 170
DB 612 GCTTTAGTCCCTGGCCCGGACATCTTGTCTCAAGTCAGAGCGCGGAGATCAGGAAC 671

QY 171 -----GlnAlaHisArgLysGlu 176

DB 672 GCAACAAGCCCTCCAGTACTGACCTGAAAACTTGTGACAAGAAGAA-CACCAACAAGTG 730
QY 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
DB 731 CTCCTGGGCTGAGGACCTTCTTCCCTCCACCCCGAAGCTGAACCTGAGGGAGAC 790
QY 197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211
DB 791 AACGGCAGAGGGAGTGCAGCAGGGGAGAAATAGCAGAGGGGCTTGG 835

RESULT 8
AAA40104
ID AAA40104 standard; cDNA; 716 BP.
AC AAA40104;
XX
DT 19-OCT-2000 (first entry)
XX
DE Human Rab10 cDNA.
XX
KW Human; Rab10; Yap/Rab family; ss.
XX
OS Homo sapiens.
XX
PN CN1249345-A.
XX
PD 05-APR-2000.
XX
PF 28-SEP-1998; 98CN-0121911.
XX
PR 28-SEP-1998; 98CN-0121911.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Tu Q, Gao J;
XX
DR WPI; 2000-400723/35.
XX
P-PSDB; AAB09979.
XX
PT Preparation of human gene coding sequence, its encoded polypeptide -
PS Claim 1; Page 17; 23pp; Chinese.
XX
CC This invention describes a novel human Rab10 cDNA sequence. The protein
CC coded by said sequence belongs to Yap/Rab family and is the homolog of
CC mouse Rab10. The present invention also relates to the polypeptide coded
CC by said nucleotide sequence and the application and preparing process
CC of said polynucleotide and said polypeptide. This sequence encodes the
CC human Rab10 protein described in the method of the invention.
XX
SQ Sequence 716 BP; 226 A; 151 C; 182 G; 157 T; 0 other;

Alignment Scores:
Pred. No.: 5.5e-52 Length: 716
Score: 530.50 Matches: 96
Percent Similarity: 82.56% Conservative: 46
Best Local Similarity: 55.81% Mismatches: 29
Query Match: 48.01% Indels: 1
DB: 21 Gaps: 1

US-09-817-198a-2 (1-212) x AAA40104 (1-716)

QY 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyVal 19
DB 78 ATGGCCAAAGACGATGACGACCTGCTTTTCAAGCTGCTCTGATCGGGGATCCGGAGTG 137

QY 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSer 39
DB 138 GGAAGACCTCGGCTCTTTTTCGTTTTTCGGATGATGCTTCAATCTACTATTATTC 197

QY 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIle 59
DB 171 -----GlnAlaHisArgLysGlu 176

PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
PS Claim 8; SEQ ID 17618; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX

SQ Sequence 3533 BP; 967 A; 716 C; 730 G; 1120 T; 0 other;

Alignment Scores:

Pred. No.:	4.85e-51	Length:	3533
Score:	530.50	Matches:	96
Percent Similarity:	82.56%	Conservative:	46
Best Local Similarity:	55.81%	Mismatches:	29
Query Match:	48.01%	Indels:	1
DB:	22	Gaps:	1

US-09-817-198A-2 (1-212) x AAH17889 (1-3533)

Qy	1	MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuLeuLeuVal	19
Db	500	ATGGCGAAGAGACGTACCGACCTGCTTTCAAGCTGCTCTGATCGGGGATTCGGAGTG	559
Qy	20	GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer	39
Db	560	GGGAAGACCTCGCTCTTTTTCGTTTTCGGATGATGCTTCAATACCTATTTATTTCC	619
Qy	40	ThrIleGlyValAspPheLeuMetLysThrIleGluValAspGlyLeuLysValArgIle	59
Db	620	ACCATAGGAATAGACTTCAAGATCAAAACAGTTGAAATACAGGAAGAAGATCAAGCTA	679
Qy	60	GlnIleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg	79
Db	680	CAGATATGGATACAGCAGCGAGCGAGGATTCACACCATCACAACCTCTACTACAGA	739
Qy	80	ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle	99
Db	740	GGCGAATGGGTATCATCTAGTATATGACATCACCAATGGTAAAGTTTGAAACATC	799
Qy	100	MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle	119
Db	800	AGCAATGGCTTAGAACATAGATGAGCATGCCAATGAAGATGGGAAGATGTACTA	859
Qy	120	GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu	139
Db	860	GGAAACAAGTGTGATATGACACACAAAAGAGTTGTACCTAAAGGAAGAGGACAGATT	919
Qy	140	AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys	159
Db	920	GCAAGGGAGCATGCTATTTAGTTTTCGAGACTAGTGCAAAAAGCAAAATATAAATCGAA	979

Qy	160	GlusPheThrArgLeuThrGluLeuValLeuGln	171
Db	980	AAGGGTTCCTCACGTTAGCTAGCATATCTCTCGA	1015

RESULT 14

ID	AA571453	standard; cDNA: 674 BP.
XX	AA571453;	
XX	13-FEB-2002	(first entry)
DE	DNA	encoding novel human diagnostic protein #7257.
KW	Human;	chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement;	medical imaging; diagnostic; genetic disorder; ss.
OS	Homo sapiens.	
PN	WO200175067-A2.	
XX	11-OCT-2001.	
XX	30-MAR-2001;	2001WO-US08631.
XX	31-MAR-2000;	2000US-0540217.
PR	23-AUG-2000;	2000US-0649167.
XX	(HYSE-) HYSEQ INC.	
PI	Drmanac RT,	Liu C, Tang YT;
DR	WPI:	2001-639362/73.
DR	P-FSDB;	ABG07266.
PT	New isolated polynucleotide and encoded polypeptides, useful in	
PT	diagnostics, forensics, gene mapping, identification of mutations	
PT	responsible for genetic disorders or other traits and to assess	
PT	biodiversity -	

Claim 1; SEQ ID No 7257; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AA564197-AA594564 represent novel human diagnostic coding sequences of the invention.
Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 674 BP; 218 A; 136 C; 173 G; 147 T; 0 other;

Alignment Scores:

Pred. No.:	6.62e-52	Length:	674
Score:	529.50	Matches:	96
Percent Similarity:	82.56%	Conservative:	46
Best Local Similarity:	55.81%	Mismatches:	29
Query Match:	47.92%	Indels:	1

Db 438 ATTGACTATGGATTAAATCTTTGGAGACAAGCGCAAAATCCAGTGCATAATGTAGAAGAG 497
Qy 161 Ser---PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLys 175
Db 498 GCATTTTTCACCTTGCACGAGATATAATGACAAAACTCAACAGAAAA 545

Search completed: November 17, 2002, 20:47:19
Job time : 260 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 20:38:29 ; Search time 20 Seconds

(without alignments)
159.645 Million cell updates/sec

Title: US-09-817-198A-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	212	10	US-09-817-198A-2
2	1092	98.8	218	10	US-09-817-198A-5
3	1077	97.5	212	10	US-09-817-198A-4
4	532	48.1	246	10	US-09-925-302-534
5	530.5	48.0	218	10	US-09-925-300-1571
6	528.5	47.8	207	10	US-09-794-257-8
7	512	46.3	190	10	US-09-822-860-5
8	497	45.0	162	10	US-09-834-765-766
9	478.5	43.3	201	10	US-09-967-736-3
10	469	42.4	201	10	US-09-967-736-8
11	467	42.3	222	10	US-09-820-003A-4
12	434.5	39.3	198	10	US-09-794-257-16
13	434.5	39.3	198	10	US-09-945-173-5
14	434.5	39.3	198	10	US-09-972-529-4
15	430	38.9	223	10	US-09-817-199A-4
16	429	38.8	223	10	US-09-817-199A-2
17	416	37.6	191	10	US-09-794-257-14
18	416	37.6	191	12	US-10-051-986-3
19	412.5	37.3	212	10	US-09-350-874-67

20	399	36.1	312	10	US-09-925-302-783
21	376	34.0	213	10	US-09-794-257-5
22	373	33.8	213	10	US-09-988-974-8
23	373	33.8	217	10	US-09-988-974-3
24	373	33.8	239	10	US-09-925-301-1077
25	370	33.5	201	10	US-09-822-860-2
26	353.5	32.0	624	10	US-09-834-765-5
27	353.5	32.0	625	10	US-09-834-765-762
28	353.5	32.0	832	10	US-09-834-765-2
29	351	31.8	168	10	US-09-834-765-765
30	343	31.0	161	10	US-09-834-765-763
31	339	30.7	208	9	US-10-108-605-45
32	338.5	30.6	216	10	US-09-945-173-10
33	338.5	30.6	217	10	US-09-925-300-1364
34	323.5	29.3	211	12	US-10-051-986-6
35	321	29.0	173	10	US-09-820-003A-2
36	320	29.0	259	12	US-10-051-986-1
37	320	29.0	260	12	US-10-051-986-4
38	310	28.1	157	10	US-09-834-765-764
39	303.5	27.5	154	10	US-09-822-860-4
40	300.5	27.2	106	10	US-09-867-550-1812
41	297	26.9	201	10	US-09-988-974-5
42	297	26.9	201	10	US-09-988-974-9
43	297	26.9	209	10	US-09-864-761-42996
44	290	26.2	144	10	US-09-972-529-7
45	282.5	25.6	208	10	US-09-925-302-629

ALIGNMENTS

RESULT 1

US-09-817-198A-2

; Sequence 2, Application US/09817198A

; Patent No. US20020146758A1

; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: CLO01188

; CURRENT APPLICATION NUMBER: US/09/817,198A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 212

; TYPE: PRT

; ORGANISM: Human

US-09-817-198A-2

Query Match 100.0%; Score 1105; DB 10; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.2e-102;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAKQYDVLFRLLIGDSGVGKTCCLRCFTDNEFHSHISTIGVDFKMKTIYVDGKVRQ	60
Db	1	MAKQYDVLFRLLIGDSGVGKTCCLRCFTDNEFHSHISTIGVDFKMKTIYVDGKVRQ	60
QY	61	IWDTAGQERYQTITTKQYRRAGQIFLVYDISSERSYQHIMKWSVDYDEYAPGVQKILIG	120
Db	61	IWDTAGQERYQTITTKQYRRAGQIFLVYDISSERSYQHIMKWSVDYDEYAPGVQKILIG	120
QY	121	NKADEQKQVREGQOQALAKYGMDFYTSACTNINIKESFRTLTELVLQHRKELEGL	180
Db	121	NKADEQKQVREGQOQALAKYGMDFYTSACTNINIKESFRTLTELVLQHRKELEGL	180
QY	181	RMRSNELALALEEEEGKPEGPANSSKTCWC	212
Db	181	RMRSNELALALEEEEGKPEGPANSSKTCWC	212

RESULT 2

US-09-817-198A-5
; Sequence 5, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Human
US-09-817-198A-5

Query Match 98.8%; Score 1092; DB 10; Length 218; 1;
Best Local Similarity 97.2%; Pred. No. 2.4e-101;
Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIIEVDGKVRQ 60

QY 61 IWDTAGQERYQTITKQYRRAQGFILVYDISSERSYQHIMKWSVDVDEYAPGVQKILIG 120
DB 61 IWDTAGQERYQTITKQYRRAQGFILVYDISSERSYQHIMKWSVDVDEYAPGVQKILIG 120

QY 121 NKADBEQKRVGRGQOQLAKEYGMDFYETSACTNLIKESFTRLTVELVLAHRKEGL 174
DB 121 NKADBEQKRVGRGQOQLAKEYGMDFYETSACTNLIKESFTRLTVELVLAHRKEGL 180

QY 175 KELEGLRMRASNELALAELEEEGKPEGPANSSKTCWC 212
DB 181 KELEGLRMRASNELALAELEEEGKPEGPANSSKTCWC 218

RESULT 3
US-09-817-198A-4
; Sequence 4, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-817-198A-4

Query Match 97.5%; Score 1077; DB 10; Length 212;
Best Local Similarity 97.6%; Pred. No. 7.1e-100;
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIIEVDGKVRQ 60

QY 61 IWDTAGQERYQTITKQYRRAQGFILVYDISSERSYQHIMKWSVDVDEYAPGVQKILIG 120
DB 61 IWDTAGQERYQTITKQYRRAQGFILVYDISSERSYQHIMKWSVDVDEYAPGVQKILIG 120

QY 121 NKADBEQKRVGRGQOQLAKEYGMDFYETSACTNLIKESFTRLTVELVLAHRKEGL 180

DB 121 NKADBEQKRVGRGQOQLAKEYGMDFYETSACTNLIKESFTRLTVELVLAHRKEGL 180

QY 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212

DB 181 RTCASNELALAELEDEGKTEGPNSSKTCWC 212

RESULT 4
US-09-925-302-534
; Sequence 534, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 534
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-534

Query Match 48.1%; Score 532; DB 10; Length 246;
Best Local Similarity 47.6%; Pred. No. 1.3e-45;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSSHISTIGVDFKMKTIIEVDGKVRQ 60
DB 44 MAKAYDHLFKLLIGDSVGKTCCLIRFADNENNYISTIGIDFKIRIVDIEGKKIKLQ 103

QY 61 IWDTAGQERYQTITKQYRRAQGFILVYDISSERSYQHIMKWSVDVDEYAPGVQKILIG 120
DB 104 VWDTAGQEREKTTTAYIRGAMGIIIVYDITDEKSFENIQNMKSIKENASAGVERLLLG 163

QY 121 NKADBEQKRVGRGQOQLAKEYGMDFYETSACTNLIKESFTRLTVELVLAHRKEGL 180
DB 164 NKCDMEAKRVQEQADKLAREHGIRFFETSASKSMNVDEAFSSLDAIL-----LKSG 217

QY 181 RMRASNELALAELEEEGKPEGPANSSKTC 210
DB 218 GRRSGN-----GNKP--PSTDLTCTC 235

RESULT 5
US-09-925-300-1571
; Sequence 1571, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1571
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1571

[illegible]

US-09-820-003A-4

Query Match 42.3%; Score 467; DB 10; Length 222;
Best Local Similarity 51.9%; Pred. No. 3.2e-39;
Matches 84; Conservative 34; Mismatches 44; Indels 0; Gaps 0;

Qy 1 MAKQYDLVFRLLIGDSVGKTCLLCRFTDNEFHSHSHTIGVDFKMTIEVDGKIVRIQ 60
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 21 MNPEYDYLKLLIGDSVGKSCLLRFPADDTYTESYISTIGVDFKIRTIELDGKTIKQ 80

Qy 61 IWDTAGOERYOTIKQYRRAGQIFLVYDIDISSERSYOHIMKWSDVDVEYAPGVQKILIG 120
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 81 IWDTAGOERFRITSSYRGAHGIIVYDVTQDSFNVKQWLQEIIRHADKDNVPIVLVGNKCDLED 140

Qy 121 NKADBEQRQVRGREGOQQLAKEYGMDFYETSACTNLNIKESF 162
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 141 NKCDLTKKVVDDYTTAKEFADSLGIPFLETSAKNATNVQESF 182

RESULT 12
US-09-794-257-16
; Sequence 16, Application US/09794257
; Patent No. US20020098041
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam accession number PF00071
US-09-794-257-16

Query Match 39.3%; Score 434.5; DB 10; Length 198;
Best Local Similarity 50.9%; Pred. No. 4.7e-36;
Matches 89; Conservative 30; Mismatches 43; Indels 13; Gaps 3;

Qy 10 RLLIGDSVGKTCLLCRFTDNEFHSHSHTIGVDFKMTIEVDGKIVRIQIWDTAGOER 69
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1 KLVIGDSVGKSSLLIRFTDNKFVEEYIPTIGVDFYTKTVEVDGKTVKLQIWDTAGOER 60

Qy 70 YQITIKQYRRAGQIFLVYDIDISSERSYOHIMKWSDVDVEYA--PEGVOKILIGNKAD--- 124
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 61 FRALRPAYRGAQGLLVYDITSRDSFENVKWLKEILRHADKDNVPIVLVGNKCDLED 120

Qy 125 -----BEQKRVGREGOQQLAKEYG--MDFYETSACTNLNIKESFTRLTELVLQ 171
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 121 DEDLELTGQKRVVSTEGEALAKELGALPFMETSAKTNTNVEEAFEELAREILK 175

RESULT 13
US-09-945-173-5
; Sequence 5, Application US/09945173
; Patent No. US20020127568A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 47324, A NOVEL HUMAN G-PROTEIN AND USES
; FILE REFERENCE: 38155-20035.00
; CURRENT APPLICATION NUMBER: US/09/945,173
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,293
; PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-945-173-5

Query Match 39.3%; Score 434.5; DB 10; Length 198;
Best Local Similarity 50.9%; Pred. No. 4.7e-36;
Matches 89; Conservative 30; Mismatches 43; Indels 13; Gaps 3;

Qy 10 RLLIGDSVGKTCLLCRFTDNEFHSHSHTIGVDFKMTIEVDGKIVRIQIWDTAGOER 69
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1 KLVIGDSVGKSSLLIRFTDNKFVEEYIPTIGVDFYTKTVEVDGKTVKLQIWDTAGOER 60

Qy 70 YQITIKQYRRAGQIFLVYDIDISSERSYOHIMKWSDVDVEYA--PEGVOKILIGNKAD--- 124
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 61 FRALRPAYRGAQGLLVYDITSRDSFENVKWLKEILRHADKDNVPIVLVGNKCDLED 120

Qy 125 -----BEQKRVGREGOQQLAKEYG--MDFYETSACTNLNIKESFTRLTELVLQ 171
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 121 DEDLELTGQKRVVSTEGEALAKELGALPFMETSAKTNTNVEEAFEELAREILK 175

RESULT 14
US-09-972-529-4
; Sequence 4, Application US/09972529
; Patent No. US20020150916A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 47316, A NOVEL HUMAN G-PROTEIN AND USES
; FILE REFERENCE: 38155-20041.00
; CURRENT APPLICATION NUMBER: US/09/972,529
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/237,716
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-972-529-4

Query Match 39.3%; Score 434.5; DB 10; Length 198;
Best Local Similarity 50.9%; Pred. No. 4.7e-36;
Matches 89; Conservative 30; Mismatches 43; Indels 13; Gaps 3;

Qy 10 RLLIGDSVGKTCLLCRFTDNEFHSHSHTIGVDFKMTIEVDGKIVRIQIWDTAGOER 69
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 1 KLVIGDSVGKSSLLIRFTDNKFVEEYIPTIGVDFYTKTVEVDGKTVKLQIWDTAGOER 60

Qy 70 YQITIKQYRRAGQIFLVYDIDISSERSYOHIMKWSDVDVEYA--PEGVOKILIGNKAD--- 124
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 61 FRALRPAYRGAQGLLVYDITSRDSFENVKWLKEILRHADKDNVPIVLVGNKCDLED 120

Qy 125 -----BEQKRVGREGOQQLAKEYG--MDFYETSACTNLNIKESFTRLTELVLQ 171
I :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 121 DEDLELTGQKRVVSTEGEALAKELGALPFMETSAKTNTNVEEAFEELAREILK 175

RESULT 15
US-09-817-199A-4
; Sequence 4, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:

